

Volume 19 Number 1

May 1968

The British Journal for the Philosophy of Science



CAMBRIDGE UNIVERSITY PRESS

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

Editors: MARY B. HESSE, D. H. MELLOR

Review Editors: B. M. FLOOD, G. M. K. HUNT

Subscriptions

The *Journal* is published quarterly (May, August, November and February). The subscription price is 60s. net in U.K. (U.S.A. \$9.50) per volume post free; single issues 20s. net in U.K. (U.S.A. \$2.75) plus postage. The subscription is 40s. for members of the British Society for the Philosophy of Science, who should order direct from the publisher. Orders may be sent to any bookseller or direct to Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1. Orders from U.S.A. and Canada should be sent to Cambridge University Press American Branch, 32 East 57th Street, New York, N.Y. 10022, U.S.A.

Notes to Contributors

Articles and correspondence should be sent to the Editors, and books for review to the Review Editors, at the Whipple Museum, Free School Lane, Cambridge. Articles should not normally exceed 5,000 words, and should be accompanied by abstracts of about 100 words which will then appear in *Dialectica*, *Synthese* and *Philosophy of Science*.

All contributions should be typed in duplicate on one side only with wide margins and double-spaced. Footnotes, which should be as few as possible, should be typed separately, double-spaced at the end of the text. They should be referred to by consecutive numbering throughout the text, in which they will be inserted in page proof. Bibliographical references should be listed alphabetically, numbered, and referred to by these numbers in the text rather than in footnotes. The style of references should follow normal *Journal* practice.

Long quotations are printed in small type without quotes, and should be so marked in the typescript. Diagrams, tables, and illustrations should be on separate sheets, with their desired position in the text clearly indicated. Symbolic formulae appearing as separate lines should be clearly distinguished in the text, and foreign letters (Greek, German, etc.) indicated in the margin. On matters of style and layout, authors should consult T. W. Chaundy et al.: *The Printing of Mathematics* (Oxford University Press, 1964).

Offprints of articles should be ordered from the Editor when corrected galley proofs are returned. 50 are supplied free; the price of extra copies is obtainable from the publisher.

Advertising

Particulars of advertising in the *Journal* may be obtained from Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1.; or in U.S.A. and Canada, Cambridge University Press American Branch, 32 East 57th Street, New York, N.Y. 10022, U.S.A.

Cik-H06105-11-270929

270929

The British Journal for the Philosophy of Science

501.05

Q771



Volume 19

CAMBRIDGE UNIVERSITY PRESS 1969

CONTENTS

ARTICLES, DISCUSSIONS, MISCELLANEOUS

ARTICLES

<i>page</i>	
1	CAN THERE BE NECESSARY CONNECTIONS BETWEEN SUCCESSIVE EVENTS? Nicholas Maxwell
27	SCIENTIFIC REVOLUTION FOR EVER? William Kneale
43	DEFINITION AND DISCOVERY (II) Fred Wilson
93	EXPLANATION AND VALUE NEUTRALITY James Leach
109	ROBOTS, CONSCIOUSNESS, AND PROGRAMMED BEHAVIOUR Keith Gunderson
123	CORROBORATION, EXPLANATION, EVOLVING PROBABILITY, SIMPLICITY AND A SHARPENED RAZOR I. J. Good
185	HIDDEN VARIABLES AND THE COPENHAGEN INTERPRETATION— A RECONCILIATION Jeffrey Bub
211	THE CASE AGAINST TELEOLOGICAL REDUCTIONISM Larry Wright
225	SOME ASPECTS OF CARNAP'S THEORY OF INDUCTIVE INFERENCE Carl-Erik Särndal
281	VAGUENESS, INEXACTNESS, AND IMPRECISION R. G. Swinburne
301	ACTION, CAUSALITY, AND TELEOLOGY Ruth Macklin
317	THE 'PRESENT' IN PHYSICS H. A. C. Dobbs
325	THE STRUCTURE OF ACCEPTANCE AND ITS EVIDENTIAL BASIS P. M. Williams

DISCUSSIONS

57	CONFIRMATION STILL WITHOUT PARADOXES William H. Baumer
63	MILLER'S PARADOX OF INFORMATION Jeffrey Bub and Michael Radner

iv *Contents*

page

- 67 LENGTH CONTRACTION AND CLOCK SYNCHRONISATION: THE
EMPIRICAL EQUIVALENCE OF THE EINSTEINIAN AND LORENT-
ZIAN THEORIES
Jon Dorling
- 70 A NOTE ON CONSILIENCE
L. Jonathan Cohen
- 145 THE STRAIGHT AND NARROW RULE OF INDUCTION:
A REPLY TO DR BUB AND MR RADNER
David Miller
- 152 THE CONCEPTUAL POSSIBILITY OF TIME TRAVEL
George Berger
- 155 HUMAN AND MACHINE LOGIC: A REJOINDER
J. R. Lucas
- 156 THE WHITE SHOE *Qua* HERRING IS PINK
I. J. Good
- 247 THE JONES CASE
William L. Harper
Henry E. Kyburg
- 251 A NOTE ON TWO 'PROBLEMS' OF INDUCTION
Paul K. Feyerabend
- 254 SIMULTANEITY AND THE FUTURE
Errol E. Harris
- 257 A NOTE ON GOODMAN'S PARADOX
J. E. J. Altham
- 345 NEW MYSTERIES FOR OLD: THE TRANSFIGURATION OF
MILLER'S PARADOX
William W. Rozeboom
- 353 NEUTRALITY, CONTINGENCY AND UNDECIDABILITY
Dharmendra Kumar
- 357 GÖDEL'S THEOREM IS A RED HERRING
I. J. Good

MISCELLANEOUS

- 88, 181,
275, 374 RECENT PUBLICATIONS ON THE PHILOSOPHY OF SCIENCE

- 85 ABSTRACTS

- 73, 159,
259, 359 REVIEWS

INDEX OF AUTHORS

ARTICLES

- BUB, JEFFREY, Hidden Variables and the Copenhagen Interpretation—A Reconciliation, 185
- DOBBS, H. A. C., The 'Present' in Physics, 317
- GOOD, I. J., Corroboration, Explanation, Evolving Probability, Simplicity and a Sharpened Razor, 123
- GUNDERSON, KEITH, Robots, Consciousness, and Programmed Behaviour, 109
- KNEALE, WILLIAM, Scientific Revolution for Ever?, 27
- LEACH, JAMES, Explanation and Value Neutrality, 93
- MACKLIN, RUTH, Action, Causality, and Teleology, 301
- MAXWELL, NICHOLAS, Can there be Necessary Connections between Successive Events?, 1
- SÄRNDAL, CARL-ERIK, Some Aspects of Carnap's Theory of Inductive Inference, 225
- SWINBURNE, R. G., Vagueness, Inexactness, and Imprecision, 281
- WILLIAMS, P. M., The Structure of Acceptance and its Evidential Basis, 325
- WILSON, FRED, Definition and Discovery (II), 43
- WRIGHT, LARRY, The Case Against Teleological Reductionism, 211

DISCUSSIONS

- ALTHAM, J. E. J., A Note on Goodman's Paradox, 257
- BAUMER, WILLIAM H., Confirmation Still Without Paradoxes, 57
- BERGER, GEORGE, The Conceptual Possibility of Time Travel, 152
- BUB, JEFFREY, and RADNER, MICHAEL, Miller's Paradox of Information, 63
- COHEN, L. JONATHAN, A Note on Consilience, 70
- DORLING, JON, Length Contraction and Clock Synchronisation: The Empirical Equivalence of the Einsteinian and Lorentzian Theories, 67
- FÉYERABEND, PAUL K., A Note on Two 'Problems' of Induction, 251
- GOOD, I. J., The White Shoe *Qua* Herring is Pink, 156
- GOOD, I. J., Gödel's Theorem is a Red Herring, 357
- HARPER, WILLIAM L. (with KYBURG, HENRY E.), The Jones Case, 247
- HARRIS, ERROL E., Simultaneity and the Future, 254
- KUMAR, DHARMENDRA, Neutrality, Contingency and Undecidability, 353
- KYBURG, HENRY E. (with HARPER, WILLIAM L.), The Jones Case, 247
- LUCAS, J. R., Human and Machine Logic: A Rejoinder, 155
- MILLER, DAVID, The Straight and Narrow Rule of Induction, A Reply to Dr Bub and Mr Radner, 145
- RADNER, MICHAEL, and BUB, JEFFREY, Miller's Paradox of Information, 63
- ROZEBOOM, WILLIAM W., New Mysteries for Old: The Transfiguration of Miller's Paradox, 345

REVIEWS

- BASRI, SAUL A. *A Deductive Theory of Space and Time*; review by David Bohm, 173
- COLLINS, N. L. and MICHIE, D., Eds. *Machine Intelligence 1*; review by Margaret A. Boden, 271
- DALE, E. and MICHIE, D., Eds. *Machine Intelligence 2*; review by Margaret A. Boden, 271
- DEBUS, ALLEN G. *The Chemical Dream of the Renaissance*; review by P. M. Rattansi, 270
- DIRENZO, GORDON J., Ed. *Concepts, Theory and Explanation in the Behavioral Sciences*; review by D. W. Hamlyn, 81
- FISCHER, R., Ed. *Interdisciplinary perspectives of time, Annals of the New York Academy of Sciences*; review by Eva Cassirer, 359
- HARTMAN, ROBERT S. *The Structure of Value: Foundations of Scientific Axiology*; review by G. R. Grice, 179
- HINTIKKA, J. and SUPPES, P., Eds. *Aspects of Inductive Logic*; review by Isaac Levi, 261
- JEFFREY, RICHARD C. *The Logic of Decision*; review by Richard Robinson, 177
- KÖRNER, STEPHAN. *Experience and Theory*; review by J. L. Mackie, 168
- LAKATOS, IMRE, Ed. *Problems in the Philosophy of Mathematics*; review by A. Sloman, 171
- LEVI, ISAAC. *Gambling with Truth*; review by J. L. Mackie, 261
- MANDELBAUM, MAURICE. *Philosophy, Science and Sense Perception*; review by D. Primack, 269
- MICHIE, D., Ed. *Machine Intelligence 3*; review by Margaret A. Boden, 271
- NEWTON, I. *The Correspondence of Isaac Newton, vol. iv. 1694-1709*. Ed. J. F. SCOTT; review by G. J. Whitrow, 268
- POPPER, KARL R. *Conjectures and Refutations*; review by Peter Achinstein, 159
- PRIOR, ARTHUR N. *Papers on Time and Tense*; review by S. Blackburn, 371
- PROKHOVNIK, S. J. *The Logic of Special Relativity*; review by J. V. Narlikar, 267
- REGNÉLL, H. *Ancient Views of the Nature of Life: Three Studies in the Philosophies of the Atomists, Plato and Aristotle*; review by J. D. G. Evans, 265
- SALMON, WESLEY C. *The Foundations of Scientific Inference*; review by Isaac Levi, 259
- SCHIEFFLER, ISRAEL. *Science and Subjectivity*; review by Mary Hesse, 176
- SCHLEGEL, RICHARD. *Completeness in Science*; review by L. Laudan, 264
- SCHMIDT, PAUL F. *Perception and Cosmology in Whitehead's Philosophy*; review by Dorothy Emmet, 83

4 FEB 1970

MEMBERS OF THE BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE *

[F = Foundation Member of the former Philosophy of Science Group
H = Honorary Member]

- Abel (P. M.), Dept of Sociology, University of Essex, Colchester, Essex
Abrams (Prof. J. W.), Committee on the History and Philosophy of Science,
621 Spadina Avenue, University of Toronto, Ontario, Canada
F Acton (Prof. H. B.), Dept of Philosophy, The University, Edinburgh
Alexander (P.), Dept of Philosophy, The University, Bristol 8
Allentuck (Prof. M.), 5 West 86 Street, Apt. 12B, New York, New York
10024, U.S.A.
Andreski (Prof. S. L.), Dept of Sociology, The University of Reading,
Whiteknights Park, Reading
F Armitage (Dr A.), University College, London, W.C.1
Ayer (Prof. A. J.), New College, Oxford
- Bailey (Miss M. L.), 28 Chester Street, London, S.W.1
Barlow (W.), 3 Albert Court, Kensington Gore, London, S.W.7
Bartley (Prof. W. W.), Dept of Philosophy, University of Pittsburgh,
Pennsylvania, U.S.A.
Bastin (Dr E. W.), 11 Millington Road, Cambridge
Beacon (G. R.), 247 Cliffe Road, Strood, Rochester, Kent
Bear (D.), 99 Castlewood Drive, Eltham, London, S.E.9
Bernal (Prof. J. D.), Birkbeck College, Malet Street, London, W.C.1
Best (Dr A. E.), 19 Farndon Road, Oxford
Blake (A. G. E.), Coombe Springs, Coombe Lane West, Kingston-upon-
Thames, Surrey
Boden (Miss M. A.), Arts II, University of Sussex, Falmer, Brighton, Sussex
Bohm (Prof. D.), Birkbeck College, Malet Street, London, W.C.1
F Boltz (C. L.), 20 Willow Road, London, N.W.3
Boon (M. R.), 6 Vicarage Road, Chester
Borger (R.), Dept of Psychology and Social Science, Brunel College, Wood-
lands Avenue, London, W.3
Bradley (J.), University of Hull, Hull, Yorks.
F Braithwaite (Prof. R. B.), King's College, Cambridge
Bramwell (D. G.), 29 Westview Avenue, Whyteleaf, Surrey
Brewer (C. V.), 16 Northampton Road, Addiscombe, Croydon, Surrey

* Revised to February 1968.

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

- Brink (D. H.), Commonwealth Hall, Cartwright Gardens, London, W.C.1
 Brodie (M.), Gaylands, High Road, Cookham Rise, Berks
 Brookes (B. C.), Dept of Electrical Engineering, University College, Gower Street, London, W.C.1
 Brookes (W. M.), 45 The Parkway, Bassett, Southampton
 Buchdahl (G.), 11 Brookside, Cambridge
 Bunge (Prof. M. A.), Dept of Philosophy, McGill University, Montreal, Canada
 Burch (P. R. J.), Dept of Medical Physics, The General Infirmary, The University, Leeds 1
- Cain (J. C.), 63 Park Road, Chiswick, London, W.4
 F Caldin (Prof. E. F.), 3 London Road, Canterbury, Kent
 Calvert (Dr H. R.), The Science Museum, London, S.W.7
 F Carter (Dr G. S.), 40 Grange Road, Cambridge
 Cassirer (Dr Eva), Dept of Logic and Metaphysics, St Salvator's College, St Andrews, Fife
 Chance (Dr M. R. A.), 12 Innage Road, Northfield, Birmingham 31
 Chilton (D.), The Science Museum, London, S.W.7
 Clark (Dr J. H.), Psychology Dept, University of Manchester, Manchester 13
 Clarke (E. J. W.), 67 Walsingham Gardens, Stoneleigh, Ewell, Surrey
 Clarke (Dr E. S.), Wellcome Historical Medical Library, Euston, London, N.W.1
 Clarke (J. D.), 4 Upper Pines, Woodmansterne, Banstead, Surrey
 Clegg (A. D.), 4 Aberdale Gardens, London, N.W.6
 Coe (Dr W. E.), Ashby Institute, Stranmillis Road, Belfast 9
 Cohen (L. J.), The Queen's College, Oxford
 Cohen (R. P. S.), London School of Economics, Houghton Street, Aldwych, London, W.C.2
 Connelly (G. H. P.), 4 Margaret Road, Bexley, Kent
 Cousin (Prof. D. R.), Dept of Philosophy, The University, Sheffield
 F Crombie (Dr A. C.), All Soul's College, Oxford
 Cuming (M. J.), 4 Strathmore Gardens, London, W.8
 Cunliffe (C. R. A.), 29 Riverdale Gardens, Twickenham, Middx
 Cunningham (A. D.), 7 Babraham Road, Cambridge
 Cutler (P.), Hawthorn Cottage, New Street, Farsley, Yorks.
- Davies (D.), 135 Aspen, Dollard des Ormeaux, P.Q., Canada
 Davies, (L. T.), Institute of Marine Science, Rickenbacker Causeway, Miami, Florida 33149, U.S.A.
 Day (J. P.), Dept of Philosophy, The University, Keele, Staffs
 Dedek (J. V. O. V.), 19 Breymisseberg, Tirlemont (Tienen), Belgium
 Demant (D. R.), 5 Grosvenor Vale, Ruialip, Middx
 Denyer (Dr G. E.), 3 Green Close, Writtle, Essex
 Dickson (J.), 153 Kingsway, Hove 3, Sussex
 HF Dingle (Prof. Herbert), 104 Downscourt Road, Purley, Surrey
 Dobbs (H. A. C.), c/o Travellers' Club, Pall Mall, S.W.1

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

- Dorling (J.), School of Mathematical and Physical Sciences, University of Sussex, Falmer, Brighton, Sussex
- Douglas (G. O.), 1 Patrick Road, West Bridgford, Nottingham
- Dresser (Dr D. W.), National Institute for Medical Research, Mill Hill, London, N.W.7
- Duckers (A. S.), Black Roof, Oak Way, Ashted, Surrey
- Duncan (A. M.), 116 Herrick Road, Loughborough, Leics
- Dutton (Miss E. A.), 78 Randolph Avenue, London, W.9
- Dyason (Diana), Dept of History and Philosophy of Science, University of Melbourne, Parkville N.2, Victoria, Australia
- Ebison (M. G.), 84 College Road, Isleworth, Middlesex
- Eccles (Sir John), Dept of Physiology, State University of New York, Buffalo, New York 14214, U.S.A.
- Edge (Dr D. O.), Science Studies Unit, Edinburgh University, 5 Buccleuch Place, Edinburgh 8
- Edge (John), 10 Dorchester Court, Herne Hill, London, S.E.24
- Edmonstone (Rev. G. R.), Dept of Mathematics, Rhodes University, Grahamstown, South Africa
- Ehrenberg (Dr A. S. C.), Aske Research Ltd, 13 Dover Street, London, W.1
- Ellis (Prof. B. D.), c/o Dept of Philosophy, Latrobe University, Bundoora, Melbourne, Victoria, Australia
- Emmett (Prof. Dorothy), 11 Millington Road, Cambridge
- Evans (Dr G. I.), 13 E. Melbury Road, London, W.14
- Evans (Dr R.), Physics Dept, University of Essex, Wivenhoe Park, Colchester, Essex
- Farrar (Dr Kathleen R.), Homeland, Navigation Road, Altrincham, Cheshire
- Fatan (Dr H. A.), 33 Clairview Road, London, S.W.16
- Feibleman (Prof. J. K.), 1424 National Bank of Commerce Building, New Orleans 12, La, U.S.A.
- Fetherstonhaugh (P. W. E.), 13 King Henry's Road, London, N.W.3
- Fincham (D. G.), 17 Eastbourne Grove, Westcliffe-on-Sea, Essex
- Findlay (A. F.), 48 Homewood Avenue, Cuffley, Potters Bar, Mddx
- Fine (Prof. A. I.), Dept of Philosophy, Cornell University, Ithaca, New York, U.S.A.
- Finlay (Miss S.), 17 Manchester Street, London, W.1
- Fischer (Prof. R.), Dept of Psychiatry, Div. Behavioral Sciences, The Ohio State University, Upham Hall, 410 West 10th Avenue, Columbus, Ohio 43210, U.S.A.
- Ford (O. E.), Flat 44, Bidston Court, Upton Road, Birkenhead, Cheshire
- Fraser (J. T.), P.O. Box 164, Pleasantville, New York 10570, U.S.A.
- Freeland (Dr G. A.), School of History and Philosophy of Science, University of New South Wales, P.O. Box 1, Kensington, N.S.W., Australia
- Fried (Dr J. M.), Pine Hill, Haslemere, nr Wycombe, Bucks
- Friedman (Miss E.), 68 Antrim Mansions, Antrim Road, London, N.W.3
- Furley (D. J.), University College, London, W.C.1

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

- Furlong (Prof. E. J. J.), Trinity College, Dublin
 Garland (J. A.), 30 Moore Street, London, S.W.3
 Giedymin (Dr J.), 53 Beaconsfield Villas, Brighton 6, Sussex
 Gillen (Dr W. H.), Marion County General Hospital, 960 Locke Street,
 Indianapolis, Indiana 46202, U.S.A.
 Gillies (D. A.), 24 Borough Lane, Saffron Walden, Essex
 Goddard (S. C.), 69 Weigh-ton Road, Harrow, Middlesex
 Good (Prof. I. J.), Dept of Statistics, College of Arts and Sciences, Virginia
 Polytechnic Institute, Blacks-bury, Virginia 24061, U.S.A.
 Goody (Dr W.), University College Hospital, Gower Street, London,
 W.C.1
 Goodstein (Prof. R. L.), The University of Leicester, Leicester
 Gorski-Popiel (J.), 92 Clova Road, Forest Gate, London, E.7
 Gort (Miss J.), Augustus Court, Tite Street, London, S.W.3
 Gower (B. S.), 50 Old Elvet, Durham
 Grant (Prof. C. K.), Dept of Philosophy, The University, Durham
 Grattan-Guinness (I. O.), 34 Hillside Gardens, Barnet, Herts
 Greenaway (F.), The Science Museum, London, S.W.7
 Grieder (A.), 10 Lawn Road, London, N.W.3
 Gross (H), 102 Princess Court, Queensway, London, W.2
 Gruner (Dr R.), 1745 Henry St., Halifax, Nova Scotia, Canada
 Hacking (Dr I. M.), Dept of Philosophy, University of British Columbia,
 Vancouver, B.C., Canada
 Halsbury (The Rt Hon The Earl of), Flat 2, 17 Stanhope Terrace, London,
 W.2
 Harper (Dr R.), 18 Leng Crescent, Bluebell Road, Norwich, Norfolk
 Harré (R.), 39 Wellington Square, Oxford
 Harris (Mrs D. G.), 10 Highbury Place, London, N.5
 Harris (F. T. C.), Middlesex Hospital Medical School, London, W.1
 Harrison (Prof. J.), 42 Broadgate, Beeston, Nottingham
 Harte (Dr E. A.), Dept of Applied Mathematics, University College of
 Swansea, Singleton Park, Swansea
 Hawton (H.), Burbank, Barnet Lane, Elstree, Herts.
 Hayek (Prof. F. A. von), 78 Freiburg I.Br., Urachstrasse 27, Germany
 Heelan (Rev Dr P. A., S.J.), Dept of Philosophy, Fordham University,
 Bronx, New York 10458, U.S.A.
 Hesse (Dr Mary B.), Whipple Science Museum, Free School Lane, Cam-
 bridge
 Higgs (Dr P. W.), Tait Institute, The University, Edinburgh, 8
 Hindmarch (Dr I.), Dept of Psychology, The University, Leeds, 2
 Hirschmann (Dr H. D.), Dept of Philosophy, University of Bristol
 Hoare (Dr M.), 20 Bradbourne Street, London, S.W.6
 Hock (Miss A. F.), 27 Heath Crescent, Stockport, Cheshire
 Hodgson (Dr P. E.), 61 The Garth, Yarnton, Oxford
 Hunt, (A. F.), 16 Maple Avenue, Leigh-on-Sea, Essex
 Hunt (G. M. K.), Kings College, Cambridge

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

- F Hutten (Dr E. H.), Garden Flat, 43 Netherhall Gardens, London, N.W.3
 Ironmonger (E. E.), Westminster College, Oxford
 Ishiguro (Miss H.), Dept of Philosophy, University College London, Gower Street, London, W.C.1
- Jackson (R. B.), The Cottage, 9c The Grove, London, N.6
 Jarvie (I. C.), Dept of Philosophy, York University, Toronto 12, Canada
 Jeffreys (Prof. Sir Harold), F.R.S., St John's College, Cambridge
 Jevons (Prof. F. R.), 'Longley', The Lowes, Bowdon, Altrincham, Cheshire
 Jones (A. M.), Penlee, Mottingham Lane, London, S.E.9
 Jordan (Dr Z. A.), Dept of Sociology, University of Reading, Whiteknights Park, Reading, Berks.
- Kamino (Keiichiro), c/o Mr H. Fukumoto, 49 Jodoji-Higashida-cho, Sakyo-ku, Kyoto, Japan
 Kapoor (Dr S. C.), University of Sussex, Falmer, Brighton, Sussex
 Katchalsky (Prof. A.), Dept of Polymer Research, The Weizmann Institute of Science, Rehovoth, Israel
 Kearney (R. A. M.), 49 Cromwell Road, Wimbledon, London, S.W.19
 Keenan (Prof. C. W.), Dept of Chemistry, University of Tennessee, Knoxville, Tennessee, 37916, U.S.A.
 Kilmister (Prof. C. W.), King's College, Strand, London, W.C.2
 King (H. C.), 'Arnon', Devonshire, Amersham, Bucks
 King (Miss P. H. M.), 37 Albion Street, Hyde Park, London, W.2
 Klein (E.), 4 St George's Drive, Westcliff-on-Sea, Essex
 Kneale (Prof. W. C.), Colton House, Burnshall, nr Skipton, Yorks
 Kneebone (Dr G. T.), Bedford College, Regent's Park, London, N.W.1
 Knox (W. C.), Sample Oak, Orchard Road, Shalford, nr Guildford, Surrey
 Koertge (Miss N.), 9A Montpelier Square, London, S.W.7
 Koestler (A.), 8 Montpelier Square, London, S.W.7
 Körner (Prof. S.), The University, Bristol 8
- Lakatos (Dr I.), London School of Economics, Houghton Street, London, W.C.2
 Lal (S.), Dept of Physiology and Pharmacology, Chelsea College of Science and Technology, Manresa Road, London, S.W.3
 Lambert (E. W.), 106 Talbot Street, Moss Side, Manchester 16
 Landau (R. S.), 10 Ladbroke Walk, London, W.11
- F Landsberg (Prof. P. T.), 13 Ely Road, Cardiff
 Laudan (L. L.), Dept of History and Philosophy of Science, University College London, Gower Street, London, W.C.1
 Lavis (D. A.), Dept of Mathematics, Chelsea College of Science and Technology, Manresa Road, London, S.W.3
 Leach (Dr E. R.), 52 Storey's Way, Cambridge
 Ledermann (Dr E. K.), 29 Avenue Road, London, N.W.8
 le Fèvre (E. J.), Queen Mary College, Mile End Road, London, E.1
 Lejewski (Prof. C.), Dept of Philosophy, University of Manchester, Manchester 13

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

- Levinson (H. B.), 42 Willoughby Road, London, N.W.3
 Lindenmayer (A.), Biology Dept, Queen's College, City University, Flushing,
 New York, 11367, U.S.A.
 Long (M. J.), Eltham College, Mottingham, London, S.E.9
 Lyon (A. J.), 24a Ravenscroft Avenue, London, N.W.11
- MacDonald-Smith (The Rev. J.), The Vicarage, Kidmore End, Reading
 Berks.
- McEvoy (J. G.), 1 Dawlish Road, London, N.W.2
 McKerrow (Dr J. C.), Albion Hotel, 4 Taverton Street, London, W.C.1
 McKay (Prof. D. M.), University of Keele, Staffs.
 Mackie (Prof. J.), University College, Oxford
 Mackie (J.), 7 Kirk Brae, Longridge, Bathgate, West Lothian
 McKinnon Wood (T. R.), The Windmill, Hildersham, Cambridge
 McLaughlin (G. H.), 12 Granville Place, High Road, Finchley, N.12
 Major (P. R.), 49 Brockman Crescent, Dymchurch, Kent
 Malcolm (Dr N. L.), Sub-Dept of Chemical Microbiology, Dept of Bio-
 chemistry, Tennis Court Road, Cambridge
 Manser (A. R.), Dept of Philosophy, The University, Southampton
 Markham (K. A.), Welsh College of Advanced Technology, Cardiff
 Marks (D. H.), 76 Westbourne Grove, London, W.2
 Marshall (Dr N.), 1 Argyll Mansions, Beaufort Street, London, S.W.3
 Martin (R. C.), 67 Tunstall Road, Croydon, Surrey
 Martins (H. G.), School of Social Studies, Dept of Sociology, University of
 Essex, Wivenhoe Park, Colchester, Essex
- F Mason (Prof S. F.), School of Chemical Sciences, Wilberforce Road, Norwich
 Maxwell (A. N.), Dept of the History and Philosophy of Science, University
 College, London, W.C.1
- Mayne (A. J.), Archers' Court, Hastings, Sussex
- F Mayo (B.), University of Birmingham, Edgbaston, Birmingham 15
 Mays (Dr W.), Dept of Philosophy, The University, Manchester 13
 Mellor (D. H.), Pembroke College, Cambridge
 Meredith (Prof. G. P.), Dept of Psychology, The University, Leeds 2
 Miller (Dr A. R.), 29 Rochester Road, Canterbury, Victoria, Australia
 Miller (D. W.), 16 Hallam Gardens, Hatch End, Harrow, Middx
 Miller (Dr E.), 77 Harley House, Marylebone Road, London, N.W.1
 Milligan (D. E.), 11 Trelawney Road, Cotham, Bristol 6
 Molland (A. G.), Dept of History and Philosophy of Science, King's College,
 Aberdeen
- Montgomery (Prof. H. A.), Dept of Philosophy, University of Auckland, Box
 2175, Auckland, New Zealand
- Morgan (B. J.), 95 Tile Kiln Lane, Bexley, Kent
- Morgenstern (N.), Dept of Civil Engineering, University of Alberta, Edmonton,
 Alberta, Canada
- Moss (J. M. B.), Dept of Philosophy, University College of Wales, Aberystwyth
- Musgrave (A. E.), London School of Economics, Houghton Street, London,
 W.C.2

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

Nidditch (Dr P. H.), Physics Building, University of Sussex, Falmer, Brighton, Sussex

North (Dr J. D.), 28 Chalfont Road, Oxford

Nowicki (Mrs M.), London School of Economics, Houghton Street, London, W.C.2

O'Connor (Prof D. J.), Dept of Philosophy, The University, Exeter

Okafor (Dr N.), Dept of Microbiology, University of Nigeria, Nsukka, Eastern Nigeria

Oldfield (Lady Kathleen), 10A Bardwell Road, Oxford

Omer (I. A.), 7 Harvey Road, Cambridge

Orr (A.), 59A Staverton Road, London, N.W.2

F Owen (Dr E. C.), The Hannah Dairy Research Institute, Kirkhill, Ayr

Ozanne (K. W.), Flat 5, Linden Court, Bradbourne Park Road, Sevenoaks, Kent

Page (Gillian, Miss), 72 Belsize Park Gardens, London, N.W.3

Parry (Mrs K. Croose), 45 Queen's Gate Gardens, London, S.W.7

Pask (A. G. S.), 20 Hill Rise, Richmond, Surrey

Peacock (Dr R. T.), Dept of Chemistry, University of Queensland, St Lucia, Brisbane, Queensland, Australia

Peacock (T. E.), 35 Wellesley Street, Jarrow, Durham

Peetz (D. W.), Dept of Philosophy, University of Nottingham, University Park, Nottingham

F Pirenne (Dr M. H.), University Laboratory of Physiology, Oxford

Pirie (Dr N. W.), F.R.S., Rothamsted Experimental Station, Harpenden, Herts

Plaut (Dr H. C.), 14 Holly Hill Drive, Banstead, Surrey

Pleydell-Pearce (A. G.), University of Birmingham: Extra-Mural Studies, P.O. Box 363, Edgbaston, Birmingham 15

Pleydell-Pearce (I. W.), 13 Southfield Road, Bristol 6

Pode (J. S. F.), c/o School Office, Eton College, Windsor, Berks

Polanyi (Prof. M.), F.R.S., 22 Upland Park Road, Oxford

Poole (H. W.), 339 Westbourne Grove, Westcliffe-on-Sea, Essex

F Popper (Prof Sir Karl R.), London School of Economics, Houghton Street, Aldwych, London, W.C.2

Post (Dr H. R.), Chelsea College of Science and Technology, Manresa Road, London, S.W.3

Post (Dr J. B.), 4 Hampstead Grove, London, N.W.3

Powers (J. H.), 118 Bartholomews, Brighton 1, Sussex

Pratt (Dr O. E.), 39 Ebbisham Road, Worcester Park, Surrey

Prokhovnik (Prof. S. J.), School of Mathematics, University of New South Wales, Kensington, N.S.W., Australia

Pursey (M.), 51 Clyst Valley Road, Clyst St. Mary, Exeter, Devon

Purton (A. C.), 20 Knatchbull Road, London, S.E.5

Quelon (F.), 133 Mottingham Road, London, S.E.9

- Ratray (C. M. I.), Dept of Mathematics (Computation), University of Manchester Institute of Science and Technology, Sackville Street, Manchester 1
- Ravest (J. F.), 18 Underhill Crescent, Abergavenny, Monmouthshire
- Ravetz (Dr and Mrs J.), 5 Moor Park Avenue, Leeds 6
- Rawcliffe (D. H.), Treliddon, Torpoint, Cornwall
- F Rawlins (F. I. G.), Danny, Hurstpierpoint, Hassocks, Sussex
- Rennie (H. G.), Dept of Economics, Room 306, Ohio State University, 1775 South College Road, Columbus, Ohio 43210, U.S.A.
- Rhodes (Michael), c/o the B.B.C., Broadcasting House, London, W.1
- J. H. Robertson, School of Chemistry, The University, Leeds
- Robinson (Prof John H.), University College, Box 9184, Dar es Salaam, Tanzania
- Robson (T. G.), Dunwood Hall, Endon, Nr. Stoke on Trent, Staffs
- Rogers (Dr Ruth H.), 3 Epworth Terrace, Aberystwyth, Cards
- Rowan-Davies (E.), 5 Bathwick Hill, Bath, Somerset
- Rudwick (Dr M. J. S.), Corpus Christi College, Cambridge
- H Russell (The Rt Hon The Earl), O.M., F.R.S., Plas Penrhyn, Penrhyn-deudraeth, Merionethshire
- F Russell (Rev. J. L.), Heythrop College, Chipping Norton, Oxford
- Saraga (Dr W.), 43 Kingsway, Petts Wood, Orpington, Kent
- Savage (Dr S. W.), Windover, Ingon Lane, Stratford-on-Avon, Warwicks
- Self (Sir Henry), 16 Vernon Terrace, Brighton, Sussex
- Serebriakoff (V.), Flat 1, No. 6, The Paragon, Blackheath, London, S.E.3
- Seton Pollock (A. F.), 76 Berkeley Avenue, Reading, Berks.
- Sewell (E. J.), North East Flat, Bede College, Durham
- Shaikh (N. A.), Suparco, P.O. Box 3125, Karachi 29, Pakistan
- Shaw (Dr R.), Dept of Mathematics, The University, Hull
- Shea (Fr W. R. G.), Darwin College, Cambridge
- Sheiham (Mrs H.), 16 Eton Road, London, N.W.3
- Shelton-Jones (E.), 82 North Road, Richmond-upon-Thames
- Shigeo (I.), Suzaka P.O. Box 8, Shin-machi, Suzaka-Shi, Nagano-Ken, Japan
- Shilton (R.), Chemistry Dept, University of Aberdeen, Aberdeen
- Siddiqui (Dr K. A.), Research Station, Research Branch, Canada Agriculture, Lethbridge, Alberta, Canada
- F Silver (Prof. R. S.), Dept of Mechanical Engineering, Heriot Watt College, Chambers Street, Edinburgh 1
- Sim (Group-Captain J. W.), Pine Cottage, Harfield Road, Sunbury on Thames, Mddx
- Simmons (H.), Dept of Mathematics, University of Bristol, Bristol 8
- Singer (Dr Jean V. L.), 54 Simons Walk, Englefield Green, Surrey
- Singer (Dr Konrad), 18 Gladstone Avenue, Whitton, Middlesex
- Sloman (Dr A.), Arts Building, University of Sussex, Falmer, Brighton
- Sluga (Dr H. D.), Dept of Philosophy, University College, London, Gower Street, London, W.C.1

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

- Smith (J. P.), 201 Morningside Road, Albert Lea, Minnesota, 56007, U.S.A.
 Smythe (M. B.), 33 Priory Road, London, N.W.6
 Smythies (Dr J. R.), University Dept of Psychiatry, D.P. Royal Edinburgh
 Hospital, Morningside Park, Edinburgh 10
 Sofronion (Dr S. A.), 21 Ferncroft Avenue, London, N.W.3
 Solntseff (Dr N.), 18 Castlenau, London, S.W.13
 StClair-Penny (A. M.), 24 Boundaries Mansions, Boundaries Road, London,
 S.W.12
 Stack-Dunne (Dr M.), M.R.C. Laboratories, Holly Hill, Hampstead, London,
 N.W.3
 Stead (The Rev. G. C.), Keble College, Oxford
 Stonier (Prof. A. W.), Makerere College, Kampala, Uganda
 Stopes-Roe (Dr H. V.), 155 Moor Green Lane, Moseley, Birmingham 15
 Straker (E. J.), 12 Soho Square, London, W.1
 Straupenicks (J.), Physics II, University of Sussex, Falmer, Brighton,
 Sussex
 Spilsbury (Dr R. I.), 2 Learmonth Terrace, Edinburgh 4
 Subotsky (M.), 14 Blomfield Court, Maida Vale, London, W.9
 Swinburne (R. G.), Dept of Philosophy, The University, Hull

 Tanner (Mrs R. C. H.), 13 Boundary Road, Wallington, Surrey
 Tarkowski (Z.), 703 Frobisher House, Dolphin Square, London, S.W.1
 Taylor (N. G.), 32 Newington Green, London, N.16
 Thompson (T. R.), 79 Mayor's Road, Altrincham, Ches.
 Thomson (A. G.), 148 Mill Lane, London, N.W.6
 Thorpe (Dr W. H.), F.R.S., Jesus College, Cambridge
 Timonen (J.), 12 Witley Court, Woburn Place, London, W.C.1
 Tinsman (J. H.), Box 457, State Teachers College, Kutztown, Pa., U.S.A.
 Toulmin (Dr S. E.), Brandeis University Waltham, Mass., U.S.A.
 Tucker (Dr J. W.), University College, Bangor, North Wales
 Turnbull (Miss A. V.), Flat 6, 52 Shepherd's Hill, Highgate, London, N.6

 van Valen (Dr Leigh), Anatomy Dept, University of Chicago, Chicago,
 Illinois 60637, U.S.A.
 Verity (Miss C. M.), 22 Princeton Street, London, W.C.1
 Vianna (Prof. M. P. M.), Caixa Postal-4020, Porto Alegre-RGS, Brazil
 Vickers (Sir Charles Geoffrey), Little Mead, Goring-on-Thames, Reading,
 Berks

 Walters (R. S.), 8 Gainsborough Road, London, N.12
 Ward (E. P.), Trio, Carlton Road, South Godstone, Surrey
 Wareing (Prof. P. F.), Dept of Botany, University College, Aberystwyth,
 Wales
 Watkins (Prof. J. W. N.), London School of Economics, Houghton Street,
 Aldwych, London, W.C.2
 Weber (B. C.), 41 Broadlands Court, Wokingham Road, Bracknell, Berks
 Webster (D.), 21 Waterpark Road, Broughton Park, Salford 7, Lancs

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

- Webster (Miss E. R.), New Hall, Huntingdon Road, Cambridge
 Wells (O. D.), Boarmans, Beaulieu, Hants
 Wetherick (Dr N. E.), Undergraduate School of Studies in Social Sciences,
 University of Bradford, Bradford 7, Yorks.
 Whitrow (Mrs A. M.), 41 Home Park Road, Wimbledon, London, S.W.19
 F Whitrow (Dr G. J.), Dept of Mathematics, Imperial College, London, S.W.7
 F Whyte (L. L.), 93 Redington Road, London, N.W.3
 F Wightman (Dr W. P. D.), F.R.S.E., 112 Rutten Lane, Yarnton, Oxon.
 Wilkie (Dr J. S.), Flat 6, 1/5 Hornton Street, London, W.8
 Williams (Prof. B. A. O.), King's College, Cambridge
 Williams (D. G. M.), Red Acre, Macclesfield Road, Alderley Edge,
 Cheshire
 Williams (J. E. M.), 26 Miskin Road, Dartford, Kent
 Williamson (Miss E. M.), Trumpeter's House, Old Palace Yard, Richmond,
 Surrey
 Winter (Dr H. J. J.), 43 Tollards Road, Countess Weir, Exeter
 Wisdom (Dr J. O.), Dept of Philosophy, York University, Toronto, Ontario,
 Canada
 F Withers (R. F. J.), Middlesex Hospital Medical School, Cleveland Street,
 London, W.1
 Wood (Dr R. H.), 290 Hempstead Road, Watford, Herts
 Woodger (Prof. J. H.), Tanhurst, Epsom Downs, Surrey
 Wright (Colin B.), Dept of Philosophy, University of Exeter, Exeter, Devon
 Wright (Dr F. I.), 49 Wellington Road, Wealdstone, Middlesex
 Young (Dr R. M.), 48 Eden Street, Cambridge
 Zimmerman (Prof. J.), Dept of Physics, University of Nebraska, Lincoln,
 Nebraska 68508, U.S.A.

Can there be Necessary Connections between Successive Events?

by NICHOLAS MAXWELL

I

The aim of this paper is to refute Hume's contention that there cannot be logically necessary connections between successive events. I intend to establish, in other words, not 'Logically necessary connections do exist between successive events', but instead the rather more modest proposition: 'It may be, it is possible, as far as we can ever know for certain, that logically necessary connections do exist between successive events'.

Towards the end of the paper I shall say something about the *implications* of rejecting Hume's contention.

It should perhaps be noted at the outset that Hume does not, in a sense, deny the *idea* of necessary connections between events: but he does deny the possibility of the objective existence of logically necessary connections between events. According to Hume the *idea* results from habit or custom after the observed frequent conjunction of two events. Thus he declares: 'Upon the whole, necessity is something that exists in the mind, not in objects. . . . Either we have no idea of necessity, or necessity is nothing but that determination of the thought to pass from causes to effects, and from effects to causes, according to their experienced union.'¹ In other words we may, if we wish, hold that Hume does not reject, but merely gives an analysis of, the notion of 'necessary connection', but in this case we must admit that his analysis excludes the possibility of the existence of any kind of logically necessary connections between events. Looked at in this way, the main task of this paper is to analyse, or interpret, 'There exists a logically necessary connection between two successive events E_1 and E_2 ', in such a way that this statement cannot be refuted on *a priori* grounds.

It must be admitted of course that it is *propositions*, and not *events*, that can be logically related or connected, and hence that there is something ill-formed about the assertion 'There exists a *logically* necessary connection between E_1 and E_2 '. In view of this I suggest that this statement is to be analysed as, or, is to be interpreted to mean:

(1) At the time of occurrence of E_1 there exists that which can only be

Received 15.viii.67

¹ D. Hume, *A Treatise of Human Nature*, vol. 1 (London, 1911), pp. 163-4.

completely described by propositions, P_1 , which logically imply propositions, P_2 , that state that E_2 occurs subsequently.

The essential requirement here is that, at the time of occurrence of E_1 , there exists something, X say, which is such that *any* propositions which completely describe X , logically imply propositions which state that E_2 occurs subsequently. No one denies that propositions, P_3 say, which *incorporate* a description of what exists at the time of occurrence of E_1 can logically imply propositions which state that E_2 occurs subsequently. But if P_3 only *incorporate* a description of what exists at the time of occurrence of E_1 , it is possible that, by whittling away at P_3 , one can arrive at propositions, P_4 say, which (a) give just as complete a description as P_3 of what exists at the time of occurrence of E_1 , (b) do *not* logically imply propositions which state that E_2 occurs subsequently. Thus the fact that propositions which *incorporate* a description of what exists at the time of occurrence of E_1 can logically imply propositions which state that E_2 occurs subsequently, does not suffice at all to establish the existence of a logically necessary connection between E_1 and E_2 in the sense of (1) above.

As far as I can see, (1) above successfully captures all that anyone could hope or want to mean by 'There exists a logically necessary connection between E_1 and E_2 '. Granted the truth of (1), that which exists at the time of occurrence of E_1 determines the subsequent occurrence of E_2 with logical necessity, in the sense at the time of occurrence of E_1 there is that which exists— X say—which is such that *any* propositions which completely describe X logically imply propositions which state that E_2 occurs subsequently. The 'logical connection' is of course between the propositions P_1 and P_2 , not between the events E_1 and E_2 ; nevertheless this 'logical connection' arises not merely as a result of the way in which we choose to think of, or describe, E_1 and E_2 ; it arises as a result of *that which exists* at the time of occurrence of E_1 . In asserting (1) one makes an *existential* claim, namely that at the time of occurrence of E_1 there *exists* that which can be completely described only by propositions which logically imply propositions which state that E_2 occurs subsequently.

I assume here that Hume, in denying the possibility of the existence of logically necessary connections between events, would wish to deny the possibility of (1) being true, for any two successive events E_1 and E_2 . In what follows I shall occasionally employ 'There exists a necessary connection between E_1 and E_2 ' as a shorthand expression for (1).

I turn now to a consideration of arguments, including Hume's, which may be deployed against the thesis that it is possible, conceivable, as far as

we can ever know for certain, that (1) is true, for some pairs of successive events E_1 and E_2 .

These arguments fall into three categories. In the first place it may be argued that (1) is *meaningless*. Secondly it may be argued that (1) is analytically false, that it is not a conceptual possibility that (1) could be true. But finally it might be argued that (1) is analytically *true*. The point is that in asserting (1) of two events, E_1 and E_2 , one is making an *existential* claim. If the thesis that it is possible that (1) is true is to amount to anything, then (1) must be interpreted in such a way that it is not true in any universe whatsoever, i.e. not analytically true.

One main argument employed by Hume can be dismissed without much discussion. Hume argues, essentially, that it cannot be meaningful to assert that there exists a necessary connection between two events E_1 and E_2 because there is no *impression* which could give rise to the *idea* of such a necessary connection. A basic tenet of Hume's Treatise is '*all our ideas are copied from our impressions*'.

In somewhat more modern terminology this argument might be presented as follows:

(a) Only those words (sentences) which can be given a phenomenalistic analysis, i.e. which can be analysed in terms of words (sentences) that refer to sense data, are meaningful.

(b) 'Necessary connection' as employed in the sentence 'There exists a necessary connection between E_1 and E_2 ' cannot be given such an analysis. Hence 'There exists a necessary connection between E_1 and E_2 ' cannot be meaningful.

I assume without discussion that no-one today would wish to defend any such phenomenalistic criterion of meaning as (a) above.

Much of Hume's discussion concerning causation may be construed as a defence of the *epistemological* thesis that it is impossible to *know with complete certainty* that any particular observed event, which we describe perhaps as E_1 , will be followed by an event that we describe as E_2 . This *epistemological* thesis I accept without reservation: we must just admit that it is always possible, as far as we can ever know for certain, that things may suddenly cease to happen as we expect or predict. However this *epistemological* thesis does not imply that the *ontological* thesis 'There exist necessary connections between successive events' (where this is understood as in (1)) is false. These two theses are in fact perfectly compatible. For suppose the ontological thesis is true. Suppose that at the time of occurrence of a particular observed event E_1 there exists that which can only be completely described by propositions P_1 which logically imply a

proposition P_2 that states that E_2 occurs subsequently. It is perfectly compatible with this to assert that at the time of occurrence of E_1 it is impossible to know for certain that E_2 will occur subsequently, for of course it may well be impossible to know for certain that P_1 is true.

I would wish in fact to argue that there is no empirical statement that can be known to be true with complete certainty. I support, in other words, Einstein's dictum: 'In so far as propositions refer to reality they are not certain: in so far as they are certain they do not refer to reality'.¹

Briefly, then, the position defended here is this: (a) we cannot be certain that E_2 , say, will follow a particular event we observe and describe perhaps as E_1 ; but equally (b) we cannot be certain that E_2 does not necessarily follow E_1 (in the sense of (1)).

Another argument, due to Hume, which may be deployed against the thesis that (1) is possibly true, runs, briefly, somewhat as follows. Given any two successive events, E_1 and E_2 , it is always possible that E_1 should not be followed by E_2 , possible in the sense that (a) we can imagine that E_1 is not followed by E_2 , or (b) no contradiction is involved in asserting that E_1 is not followed by E_2 . (In what follows the first criterion is ignored, since we are here concerned with *logical* necessity.) Since it is logically possible that E_1 is not followed by E_2 , it cannot be logically necessary that E_1 is followed by E_2 .

But is it always logically possible, given any two successive events E_1 and E_2 , that E_1 should not be followed by E_2 ? Is the statement ' E_1 is not followed by E_2 ' invariably non-contradictory? Suppose E_1 , that which exists at the instant in question, is such that any complete description, or specification, of E_1 logically implies propositions which state that E_2 occurs subsequently. Suppose that the meaning of ' E_1 ' is such that the statement ' E_1 occurs' logically implies some such statement as ' E_2 occurs subsequently'. In this case ' E_1 is not followed by E_2 ' is a logical contradiction. Of course no doubt if one gives only a certain kind of true description of the two events—a description in terms of what the two events look like, for example—so that E_1 is described as ' E_1' ', E_2 as ' E_2' ', let us say, then it may well be that ' E_1' is not followed by E_2' ' is invariably non-contradictory. But *this* result does not establish the inconceivability of (1).

Clearly in order to show that (1) cannot possibly be true it is necessary to show that that which exists at any instant *cannot conceivably be such that* only propositions which logically imply some proposition which states that such and such occurs subsequently, can completely state what it is that exists at the instant in question.

¹ Einstein's remark referred specifically to *mathematical* propositions: see A. Einstein, *Geometrie und Erfahrung*, Springer (Berlin, 1921), pp. 3-4.

This thesis may be defended as follows. Given any two successive events, E_1 and E_2 , if a proposition, P_1 , which is intended to be a description of E_1 , logically implies that E_2 follows E_1 , then ' P_1 ' must include some such statement as 'This event is followed by E_2 ' or ' E_1 is followed by E_2 '. But ' E_1 is followed by E_2 ' cannot be interpreted as referring exclusively to what exists at the time of occurrence of E_1 . ' E_2 follows E_1 ' refers to a *relational* property of E_1 ; it can no more be said to refer exclusively to what exists at the time of occurrence of E_1 than say 'A is to the left of B' can be said to refer exclusively to what exists at the place where A is. Hence (1) cannot possibly be true of any events.

My reply to this argument is as follows. There are at least two distinct ways in which 'event', 'object' and 'change' may be conceived, two conceptual schemes, which I shall call C_1 and C_2 . The above 'relational' argument against the possibility of (1) being true, is valid granted C_1 , invalid granted C_2 .

I shall give here only a brief, informal description of C_1 and C_2 .

C_1 and C_2 may be contrasted as follows. C_1 is committed in the first instance to an ontology of 'events' or 'time-slices', statements about objects that persist and change being analysable in terms of, or reducible to, statements about successive 'events' or 'time-slices'. C_2 , on the other hand, is committed in the first instance to an ontology of persisting, changing objects, statements about instantaneous events being reducible to statements about persisting, changing objects.

According to C_1 the universe is a sort of four dimensional array of point-events. The ultimate entities of C_1 are point-events, and the ultimate properties are properties that can be ascribed to point-events. Just as there are *spatial* relations between point-events, such as ' x is above y ', so there are *temporal* relations between point-events, such as ' x is earlier than y '. (Of course both 'above' and 'earlier than' presuppose conventionally assigned 'directions', which may however be suggested by certain pervasive features of the array of point-events.) Just as spatial relations are the essence of Space, so temporal relations are the essence of Time.

Notions of 'object', 'persistence' and 'change' can be defined in terms of the above as follows. A homogeneous spatio-temporal array of point-events can be termed an 'object': objects are thus four dimensional entities. If the object has a temporal length of t units, then the object 'persists' for those t units. If the point-events at one 'temporal end' of the object differ in a systematic way from the point-events at the other 'temporal end', then the object 'changes' in the interval t . At different moments or instants throughout t , there exist different temporal parts of the object just as different spatial parts exist at different places.

Granted this conceptual scheme one can I think legitimately argue that just as the spatial relation 'B is to the left of A' can scarcely be said to exist exclusively at the place where A exists, so the temporal relation 'E₂ is later than E₁' can scarcely be said to exist exclusively at the time of occurrence of E₁.

According to C₂ notions of 'persistence' and 'change' are of the essence of time, and are not derivable notions. Each presupposes the other. If something persists then something must change, for if there was no change, nothing would happen, there would be no 'passage of time', and hence no persistence. Again, if a thing changes then that thing must persist throughout the change. And even if it is simply the case that there are changes, without there being any one thing which suffers these changes, nevertheless a place must be identified throughout the changes, and this requires the persistence of *something*.

The ultimate entities of C₂ are objects which, by definition, persist and change. Ultimate properties of C₂, which are ascribable to objects, *presuppose the notions of persistence and change*. A description of an instantaneous event states what is true of *objects* at that instant; such a description thus presupposes the notions of persistence and change. In other words, whereas in C₁, statements about persisting, changing objects presuppose statements about successive instantaneous events that do not involve the notions of 'persistence' and 'change', in C₂, *statements about instantaneous events presuppose statements about persisting changing objects*.

According to C₂, objects have spatial parts, but no temporal parts. Objects are three dimensional, not four dimensional, entities. Certainly objects persist and change; but to say this is not to say that objects are composed of their successive instantaneous states. It is not an object, but rather the *life-history* of an object that is composed of the successive states of that object. And an object is not to be identified with its life-history.

It is, I think, clear that the ordinary common sense way of conceiving the world presupposes C₂ rather than C₁. Physical things—tables, stones, birds, etc.—are clearly 'objects' in the sense of C₂ rather than C₁. 'Object' in the sense of C₁ we should ordinarily translate as 'life-history of an object'. Instead of thinking of an object as a spatio-temporal array of events (which presupposes C₁), we ordinarily think of an event as the instantaneous state of a number of persisting, changing objects (which presupposes C₂).

We tend however to think of 'Time' in spatial terms. This is perfectly compatible with C₂ if it involves no more than simply thinking of persistence/change, i.e. time, in *metrical* terms. But if it involves thinking of objects that persist and change as being *extended* in time, as being four

dimensional entities made up of temporal parts, then clearly thinking of time in spatial terms is completely inconsistent with C_2 . The fact that we do tend to 'spatialize' time in this second sense, explains much of our ordinary puzzlement about the nature of Time. Any picture of Time which has the Present moving up an extended time-line constitutes a hopeless, contradictory fusion of C_1 and C_2 .

If one accepts C_2 , the above 'relational' argument against the possible existence of necessary connections between successive events, or states of affairs, is no longer valid, since the analogy between temporal and spatial relations collapses. I assume here that the world can legitimately be conceived in terms of C_2 , and that physical theories can be formulated in terms of C_2 .¹

It may be objected: But if there is to be any change in the world, objects must have different properties, or be in different states, at different times. Now suppose an object, X , is in the state S_1 at time t , S_2 at time $t+dt$. Then any description of what exists at time t , which logically implies ' X is S_2 at time $t+dt$ ', must incorporate some such statement as ' X is ϕ at time t ', where the *meaning* of ' ϕ ' is such that an object is ϕ at time t only if it is S_1 at time t and S_2 at time $t+dt$. But here ϕ -ness is not the sort of property that can conceivably exist at an instant, that an object can conceivably have at an instant. That which exists at time t , and is described by ' X is ϕ at time t ', is also described by ' X is S_1 at time t '. And this latter description does not logically imply ' X is S_2 at time $t+dt$ '. Hence it is not the case that that which exists at time t is such that any description of this logically implies ' X is S_2 at time $t+dt$ '. In other words, *even if C_2 is presupposed*, (1) above cannot conceivably be true.

My answer to this argument is as follows. The above argument assumes that any statements of the form ' X is ϕ at time t ', where these entail some statement about how X is, or is not, *changing* at time t , can invariably be reduced to a set of statements of the form ' X is S_0 at time t ', ' X is S_1 at time t_1 ', ' X is S_2 at time t_2 ' . . . where $t < t_1 < t_2 < . . .$, and where not one of these statements, taken on its own, entails any statement about how X is, or is not, changing at the instant in question. The above argument assumes, in other words (to state the thing a little more generally), that any statement about an object that involves the notions of 'change' or 'persistence', can be reduced to, or translated into, a set of statements that attribute similar and dissimilar properties to the object at different times, where none of these statements, taken individually, involves the notions

¹ For a defence of what is, in effect, the thesis that the special theory of relativity is compatible with C_2 , see W. Sellars, 'Time and the World Order', in H. Feigl and G. Maxwell, eds., *Scientific Explanation, Space and Time, Minnesota Studies in the Philosophy of Science*, vol. 3 (University of Minnesota Press, Minneapolis, 1962).

of 'change' or 'persistence'. *But it is precisely this assumption which must be rejected if C_2 is accepted or presupposed.* Within C_2 , there are predicates applicable to objects which presuppose the notions of 'persistence' and 'change', and which cannot be reduced to predicates which do not presuppose these notions. The thesis that there are such predicates is absolutely essential to C_2 . For the basic entity of C_2 is the 'object', which, by definition, 'persists' and 'changes'; and to say this is to say that to any object of C_2 a predicate can be applied which presupposes the notions of 'persistence' and 'change', and which cannot be reduced to predicates which do not presuppose these two notions.

At this point one might almost be tempted to conclude: If C_2 is incompatible with the above assumption, with the assumption, that is, that any predicates involving the notions of persistence and change can be reduced to predicates that do not involve these notions, then *so much the worse for C_2* . However, in the next section I shall defend a thesis, which I shall call 'dispositional realism', which entails that the above assumption is *false*.

One last remark about the distinction between C_1 and C_2 . It has sometimes been suggested that the question of whether one accepts a basic ontology of 'events', or of 'changing persisting things' (i.e. whether one accepts C_1 or C_2), depends on, or is related to, the question of whether one's basic language is *tenseless* or *tensed*.¹ The two questions are however *entirely* distinct. How one answers the question: Are propositions basically *tensed* or *tenseless*? depends simply on what one chooses to mean by the term 'proposition'. If one chooses to employ the term 'proposition' in such a way that a proposition can, with the passage of time, *change its truth-value*, then *tense* will indeed be basic and uneliminatable. According to this use of 'proposition', 'Harold Wilson is talking', for example, expresses, at different times, the *same* proposition: thus one and the same proposition can be true at one time, false at some other time. Attempts to eliminate the tense of 'Harold Wilson is talking' must fail, for such attempts will produce sentences of the form 'Harold Wilson is (tenseless) talking at time *t*'; and this sentence expresses *different propositions at different times*, granted the above interpretation of the term 'proposition'.²

One may, however, of course employ the term 'proposition' in such a way that propositions do not change their truth-value with the passage of time. In this case the sentence 'Harold Wilson is (tenseless) talking at time *t*' expresses *the same* proposition at all times, whereas 'Harold Wilson

¹ This suggestion has been defended at length by W. Sellars, see *ibid*.

² For recent developments in 'tense logic' (which is based on the idea that a proposition can have different truth-values at different times), see A. Prior, *Past Present and Future*, Oxford University Press, 1967.

is (tensed) talking' expresses *different* propositions at different times. In this case tense can be eliminated simply by reformulating any tensed sentence so that the time to which the sentence refers does not depend on the time at which the sentence is uttered.

All this has nothing whatsoever to do with whether one accepts C_1 or C_2 . For one accepts C_1 if and only if one accepts that all predicates which presuppose the notions of persistence and change can be reduced to predicates which do not presuppose these notions. And one accepts C_2 if and only if one rejects the above thesis.

The *temptation* to associate C_2 and C_1 with the above two theses about tensed propositions might arise in the following kind of way. On the one hand it is tempting to think that if the world consists, in the first instance, of persisting, changing objects, then, in the first instance, one's propositions must also persist and change, i.e. must change their time-reference and truth-value with the passage of time. But a consequence of employing the term 'proposition' in this kind of way is that tense cannot be eliminated from propositions. Thus C_2 becomes associated (for wholly fallacious reasons) with the thesis that tense cannot be eliminated. On the other hand it is also tempting to think that if one employs the term 'proposition' in such a way that propositions do not change their time-reference and truth-value with the passage of time (so that tense *can* be eliminated), then one is committed in the first instance to an ontology of 'events', or 'time-slices'. In this way the thesis that tenses can be eliminated becomes associated with C_1 .

II

So far I have been arguing against objections to the thesis that it is conceivable that (1) is true, for some two events E_1 and E_2 . In a sense, therefore, I have not yet said anything directly in support of the conceivability of (1). There remains the problem: Is it conceivable that that which exists at some instant is such that it can only be completely described by propositions which logically imply that some event occurs subsequently? In order to answer this question it will be convenient to examine whether or not the kind of descriptions we ordinarily give of things, and the descriptions of physics, could ever logically imply descriptions of subsequent states of affairs. Ultimately I hope to establish the conceivability of (1) by establishing that:

(a) It is conceivable that a *certain type* of realistic, comprehensive, deterministic physical theory is true.

(b) If a theory of this type is true, then (1) is true for successive states of affairs E_1 and E_2 .

Although the ordinary descriptions we give of things do not logically imply descriptions of subsequent states of affairs, nevertheless such descriptions do carry implications about how the things described *change*. In describing something as 'solid', 'inflammable', 'soluble', 'magnetic', 'elastic', 'hot', 'heavy', 'soft', 'sticky', or 'brittle', for example, we not only attribute a property to that thing, but we also imply that the thing will, in certain circumstances, 'behave' or 'change' in certain definite ways. This applies also to such descriptions as 'stone', 'copper', 'water', 'salt', 'glass', 'petrol', 'oxygen', etc., in that such substances have certain properties, of the above type, *by definition*.

The point to note is this: It is not just a contingent matter of fact that, for example, solid objects are objects which do not pass effortlessly through each other. It is implicit in the *meaning* of the term 'solid' that one solid object does not pass effortlessly through another solid object. Similarly it is implicit in the *meaning* of the term 'inflammable' that an object which is inflammable and which is exposed to a naked flame, itself bursts into flames. In other words, the meaning of the term 'inflammable' is such that 'X is inflammable at time *t*' and 'X is exposed to a naked flame at time *t*', taken together, logically imply 'X bursts into flame at time *t*'. Similar considerations arise in connection with the meaning of the other terms mentioned above, 'soluble', 'magnetic', 'elastic', etc.

There is, however, a very simple reason why statements of the above type, which attribute the above type of properties to objects at some instant, cannot logically imply statements which attribute properties to the objects at some subsequent instant. The reason is this: It is always a logical possibility that any object which possesses any property of the above type might suddenly and inexplicably cease to possess this property. A liquid that is inflammable at one instant might suddenly cease to be inflammable at a subsequent instant: a solid object might suddenly melt, evaporate, explode or vanish.

Of course in giving the above kind of descriptions of things we ordinarily presuppose that objects will not inexplicably vanish, petrol inexplicably cease to be inflammable, etc.; we presuppose that, in general, the above kind of properties will change only in the customary known ways. In fact a sort of general inter-related stability in the way in which properties persist and change is essential if the kind of descriptions we ordinarily give of things is to have any application at all. Inflammability, solidity, elasticity, etc., constitute properties that objects either do or do not possess only in the context of the hypothesis that these properties in general change and persist only in certain constant, inter-related ways. It is only because it is presupposed that in general an object only becomes or ceases

to be inflammable under certain specific circumstances (e.g. when soaked in petrol or water) that an assertion such as 'X is inflammable at time t ', where X is not exposed to a naked flame at time t , qualifies as a description of what exists at that instant. If nothing can be presupposed about the way in which the inflammability of an object does and does not change, then it would make, in a sense, no difference whatsoever whether one ascribed inflammability or non-inflammability to an object at a moment when it was not exposed to a naked flame, for whatever happened subsequently when the object was exposed to a naked flame, one could always argue that the object had become, remained, or ceased to be inflammable (as the case might be) in the intervening time.

The fact remains, however, that there is no precise, exceptionless, comprehensive hypothesis about the way in which such properties as inflammability, solidity, elasticity, etc., do and do not change incorporated explicitly in the meaning of the words 'inflammable', 'solid', 'elastic', etc. It is essentially for this reason that the ordinary descriptions we give of things cannot logically imply descriptions of subsequent states of affairs.

Nevertheless, the following at least can be maintained: The ordinary descriptions we give of things, the ordinary, common sense way in which we conceive the world, presuppose that:

(2) There are irreducible properties, possessed by objects at an instant, which are such that any statement which attributes such a property to an object at a time t , will, when combined with some further statement which merely states something about the state of affairs at time t , logically imply a statement about how the object is *changing* at time t .

I should perhaps add that this does not mean that *any* attribution of a 'disposition' to something constitutes ascribing a *property* to that thing. In saying of someone that he is a cigarette-smoker, for example, one says just that the person smokes cigarettes from time to time. It is not logically possible for a person to be a cigarette-smoker and never smoke a cigarette. 'X is inflammable', on the other hand, does not mean 'X bursts into flames from time to time'. An object may be inflammable even though it never bursts into flames.

I must now consider two possible objections to the claim that the ordinary descriptions we give of things presuppose thesis (2), or 'dispositional realism' as this thesis might be called.

In the first place it may be objected: An object or substance which is inflammable may not burst into flames when exposed to a naked flame for any number of different reasons. There may not be enough oxygen present; the object may be slightly damp; the flame may not be hot enough; etc.

Hence the meaning of 'inflammable' cannot be such that 'X is inflammable at time t ' and 'X is exposed to a naked flame at time t ' *logically* imply 'X bursts into flames at time t '. Similar considerations arise in connection with all the other 'dispositional properties'—solidity, elasticity, etc.

But this argument simply exploits a certain *vagueness* in the meaning of 'inflammable'. In saying that an object is inflammable one might mean anything from 'If this object is heated to 6,000°C in an atmosphere of pure oxygen, after an hour or so the object will show slight signs of oxidation', to 'The object ignites spontaneously at room temperature when exposed to the air'. In other words, the above objection no longer applies once a *precise* meaning has been assigned to 'inflammable'.

In the second place it may be objected: To ascribe inflammability or solidity, etc., to an object at some instant is not to ascribe a property to the object at that instant. Objects burst into flames or fail to ignite: but at those moments when they are not exposed to naked flames, objects do not possess some mysterious additional property of inflammability or non-inflammability. To assert 'X was inflammable at time t ', assuming X was not exposed to a naked flame at time t , is to assert no more than 'If X had been exposed to a naked flame at time t , X would have burst into flames'. In this manner any reference to a property such as inflammability, solidity, elasticity, etc., can be eliminated. Thus despite appearances to the contrary, the statement 'X is inflammable' does not *really* attribute any property to X.

I shall now attempt to refute this thesis.

The first point to note is that the thesis leaves utterly obscure what *is* to count as an ascription of a property to an object at an instant. For presumably any statement that 'supports a counter-factual conditional', i.e. implies a statement of the form 'If such and such were the case, then such and such would happen', cannot, if the above thesis is accepted, be construed as a straightforward description of what exists. But any ordinary description of things, events, states of affairs, in the world, *will* support 'counter-factual conditionals'. Bereft of such statements as 'X is solid', 'X is heavy', 'X is hot', 'X is fluid', 'X is soft', etc., and 'X consists of water', 'X consists of copper', etc., one is left entirely in the dark as to what precisely *can* be attributed to objects at any given instant.

Here then is a major objection to the thesis that reference to dispositional properties can be eliminated. Interpreting 'X is inflammable at time t ' to mean 'If X were exposed to a naked flame at time t , X would burst into flames at time t ', does not help to eliminate reference to dispositional properties, because 'is exposed to a naked flame' and 'bursts into flames' are just as dispositional as 'is inflammable'. For example, the sudden

appearance of a wavering light about an object does not mean that that object has burst into flames if the object is in no way changed or consumed with the passage of time.¹

One obvious reply to this objection is that only those statements which ascribe properties to *sense-data* can qualify as irreducible descriptions of what exists at an instant. But this thesis is equivalent to phenomenalism, and there are well known objections to phenomenalism, which I shall not discuss here.²

Granted that phenomenalism is untenable, the thesis that the ordinary descriptions we give of things do not presuppose (2) must it seems be rejected.

But in any case, what possible reason can there be for holding that statements that incorporate dispositional terms should be reducible to statements that incorporate only non-dispositional terms? Why should it be thought desirable to be able to analyse any dispositional terms such as 'flammable' in terms of purely non-dispositional terms?

One answer to this question might of course be the following: The basic descriptive terms should be such that it is always possible to verify or falsify conclusively whether a particular description is true of a particular object. But any description that incorporates a dispositional term cannot always be conclusively verified or falsified. 'X is inflammable', or 'X is solid', for example, cannot be conclusively verified or falsified at those times when X is not being tested for inflammability.

The simple answer to this is, of course, that *no* descriptions, however 'epistemologically basic', can be conclusively verified or falsified. Even on those occasions when X is being tested for inflammability or solidity, one still cannot conclusively verify or falsify 'X is inflammable' or 'X is solid', for this would require the conclusive verification of certain other descriptive statements which would in turn incorporate dispositional terms. It is only when certain low-level assumptions have been made about how relevant dispositional properties are correlated and conserved—assumptions presupposed by such 'basic' statements as 'X is a material object', 'X is water', 'X is copper', etc.—that any singular description such as 'X is inflammable at time *t*' can be verified or falsified. Assumptions about how dispositional properties are correlated and conserved can of course be tested individually: they cannot however be tested *en masse*, for then the whole

¹ That ordinary descriptive predicates are *dispositional* predicates, and presuppose the truth of laws or theories, has been repeatedly emphasised by Popper: see, for example, K. Popper, *The Logic of Scientific Discovery* (Hutchinson, London, 1959), pp. 422-6.

² For criticisms of phenomenalism see, for example, J. J. C. Smart, *Philosophy and Scientific Realism* (Routledge and Kegan Paul, London, 1963), ch. 2; and W. Sellars, *Science, Perception and Reality* (Routledge and Kegan Paul, 1963), ch. 3.

machinery of testing would break down. No observation at one instant would have any relevance to any observation at any other instant, and the fully fledged Humean nightmare would be upon one.

Of course if one wishes to defend phenomenalism, if, in other words, one accepts a sort of updated version of Hume's thesis 'All our ideas are copied from our impressions', then the meaning of terms such as 'inflammable' will seem wholly obscure unless they can be defined phenomenally, and hence non-dispositionally. But once phenomenalism, and the above criterion of meaning, have been rejected, there seems to be no very good reason why it should be thought to be *desirable* to define dispositional predicates in terms of non-dispositional predicates.

There is perhaps one other reason why it might be felt that dispositional predicates require analysis in terms of non-dispositional predicates if they are to have any clear meaning. Dispositional predicates clearly presuppose the notions of persistence and change. From statements that attribute dispositional properties to objects at some time *t* one can deduce statements about how some object is *changing* at that instant. Further, it is only in the context of assumptions about how dispositional properties are correlated and conserved—i.e. in the context of assumptions about how dispositional properties persist and change—that it becomes meaningful to attribute dispositional properties to objects at all. It may be assumed however that any statement which ostensibly attributes a property to an object at some instant, but which presupposes the notions of persistence and change, ought to be analysable in terms of statements which attribute similar and dissimilar properties to the object at successive instants, where none of these statements, taken individually, involves the notions of persistence and change. But, as I remarked in the previous section, *it is precisely this assumption which must be rejected if C_2 is accepted*. In other words, if C_2 is accepted, one has no cause whatsoever to be disturbed by the fact that the ultimate properties of things are *dispositional* properties.

Despite the undesirability and impossibility of reducing dispositional terms to non-dispositional terms, various philosophers have persisted in the attempt. And as a result, what requires analysis is not so much the meaning of 'inflammable' etc. but rather the meaning of terms coined by philosophers in their attempts to analyse 'inflammable' etc.

In the first place, the whole 'problem' of 'subjunctive conditionals' or 'contrary-to-fact conditionals' or 'counter-factual conditionals' seems to have arisen simply as a result of a refusal to countenance the idea that, for example, 'X is inflammable at time *t*' might ascribe a property to X at time *t*, where the meaning of 'inflammable' is such that the above statement, together with 'X is exposed to a naked flame at time *t*', logically implies

'X bursts into flames at time t '. For if one insists that 'X was inflammable at time t ' must be interpreted as 'If X had been exposed to a naked flame at time t , X would have burst into flames at time t ', one then creates the problem of how this *latter* statement is to be analysed.

Some philosophers have argued that this so-called 'problem of counterfactual conditionals' is to be solved along the following lines: a counterfactual conditional makes an implicit appeal to a universal conditional, where this is to be understood as asserting 'nomic' as opposed to 'accidental' universality. According to this kind of view a statement asserts 'nomic' universality if it is a universal statement that meets certain 'logical and epistemic requirements'.¹

But according to the position developed here, this sort of approach puts the cart before the horse. For, according to the view defended here, the meaning of a counterfactual conditional such as 'If this object had been exposed to a flame at time t , it would have burst into flames', at least under one legitimate interpretation, *poses no problem whatsoever*; the statement means simply 'This object was inflammable at time t (and was not exposed to a naked flame at time t)'. Hence, if the terms 'nomic universality' and 'counterfactual conditional' are linked analytically, if in other words 'A statement that asserts nomic universality supports counterfactual conditionals' is *analytic*, then 'nomic universality' should be analysed in terms of 'counterfactual conditional' and not the other way round. This can be done very simply as follows.

Consider for example the following statement:

(3) All the screws in Smith's car are rusty.

If this statement is interpreted as asserting 'accidental' universality only, it can be written as :

(4) $(x)((x \text{ is a screw} \ \& \ x \text{ is in Smith's car}) \supset x \text{ is rusty})$.

If (3) is interpreted as asserting 'nomic' universality, then it can be written as:

(5) $(x)(x \text{ is a screw} \supset x \text{ is } R)$

where 'R' is a predicate such that ' x is R at time t ' and ' x is in Smith's car at time t ' together logically imply ' x is rusty at time t '. (Here R-ability is a property analogous to inflammability or rustability: 'R-able' and 'rust-able' differ of course because, roughly speaking, it is exposure to Smith's car, not to air and water, that rusts an R-able object.)

It is I hope clear that (5) asserts 'nomic' universality because it supports the counterfactual 'If this screw were in Smith's car at time t it would be

¹ For this kind of quasi-Humean approach, see E. Nagel, *The Structure of Science* (Routledge and Kegan Paul, London, 1961), pp. 56-73.

rusty at time t' , where this means simply 'This screw is R'. (4) on the other hand asserts only 'accidental' universality because it does not support this counterfactual conditional. Any assertion of (3) would of course be interpreted as asserting only 'accidental' universality, simply because no one wishes to attribute the property R to screws.

This interpretation of the term 'nomic' universality is not put forward as the only conceivable interpretation: it is put forward simply as a legitimate interpretation, granted that there is the above analytic tie up between 'counterfactual conditional' and 'nomic universality'. In particular it must be stressed that the phrase 'asserts nomic universality', as employed here, is not intended to be equivalent to 'is a law-like statement', or 'asserts a physical law'. Clearly a statement that asserts nomic universality must satisfy further conditions before it can be called a physical law: one might, for example, stipulate that a physical law can contain no reference, explicit or implicit, to a particular object or a particular spatio-temporal region.

It may, however, be held that any reasonable interpretation of 'nomic' universality must be such that asserting nomic universality constitutes, if not a sufficient, then at least a *necessary* condition for a statement to be a physical law. And, it may be held, there are many universal statements which we wish to regard as statements of physical laws, which do not assert 'nomic' universality in the above sense.

Consider, for example, Newton's law of gravitation. If this asserts 'nomic' universality then it supports the following kind of counterfactual conditional:

(6) If this object had been over there during the interval Δt , then it would have moved in accordance with Newton's law.

But this statement, according to the above interpretation, is equivalent to attributing a dispositional property—gravitational charge—to the body during the interval Δt in question. Hence Newton's law of gravitation can only be interpreted as asserting 'nomic' universality if it is interpreted as attributing the dispositional property of gravitational charge (proportional to mass) to all bodies. But this interpretation amounts to the kind of 'essentialist' interpretation defended for example by Cotes, but rejected by Newton himself. And although Newton rejected this interpretation, he clearly did not regard the law of gravitation as asserting 'accidental' universality only. How then can one interpret a law as asserting 'nomic' universality, in the above sense, if the law is not given an 'essentialist' interpretation, i.e. if it is not interpreted as ascribing a dispositional property to objects?

The answer to this question is I suggest as follows: a law which is not, for one reason or another, given the above type of 'essentialist'¹ interpretation, may nevertheless justifiably be interpreted as asserting 'nomic' universality, in the above sense, as long as it is held to be conceivable that there is an essentialist law which entails or replaces the first law. Newton, for example, did not reject essentialism as such; he merely rejected as unacceptable the dispositional property of gravitational charge. Any theory which entailed or replaced the law of gravity, and which attributed dispositional properties to corpuscles which did not involve the notion of 'action at a distance', and which supported the counterfactual (6), would have been, it may be presumed, entirely acceptable to Newton.

To this it may be objected: But to insist that physical laws should assert 'nomic' universality in this sense is to commit physics to 'essentialism' in an important sense, even if not all acceptable physical laws are required to be essentialistic. But there are overwhelming objections to committing physics to 'essentialism' in this way. Hence physical laws cannot be required to assert 'nomic' universality in the above sense.

In reply to this objection I shall make just two comments.

(a) If there are overwhelming objections to committing physics to 'essentialism', then what the above discussion has shown is that in a perfectly straightforward, entirely legitimate, sense of 'nomic universality', physical laws do not assert nomic universality.²

(b) However, in the next section I shall in effect argue that there are no general objections to committing physics to an essentialistic programme.

III

In the last section I argued that statements that attribute dispositional properties to objects at some time *t* can logically imply statements about how some object is *changing* at time *t*. Further, I argued that it is only in the context of assumptions about how such dispositional properties are correlated and conserved—i.e. in the context of assumptions about how such dispositional properties persist and change—that it becomes meaningful to attribute dispositional properties to objects at all. It was maintained however that ordinary descriptions of objects at some instant cannot logically imply descriptions of subsequent states of affairs because there are no

¹ By 'essentialist' interpretation of a law I mean only that the law is interpreted as attributing a dispositional property—such as gravitational charge—to objects of a certain kind.

² It might be noted that if there is a perfectly straightforward (but non-Humean) sense in which only 'essentialist' laws can assert 'nomic' universality, there is also a perfectly straightforward (but non-Humean) sense in which only 'essentialist' laws can qualify as *explanatory* laws.

precise, comprehensive, exceptionless assumptions about how dispositional properties persist and change incorporated explicitly in the meaning of the terms which refer to these properties. Thus statements which ascribe inflammability, solidity, etc., to objects cannot logically imply statements about subsequent states of affairs because it is always *logically possible* that the objects might suddenly and inexplicably cease to be inflammable, solid, etc.

This suggests that in order to arrive at propositions which both qualify as descriptions of what exists at one instant and logically imply descriptions of what exists at subsequent instants, it will be sufficient to formulate a precise, comprehensive hypothesis about how things do and do not change, a comprehensive deterministic physical theory, in other words, which is explicitly incorporated in, or presupposed by, the meaning of the terms of the hypothesis in question.

More precisely, I suggest that in order to establish the conceivability of (1), it will be sufficient to establish that it is conceivable that that which exists at some instant is such that only propositions which incorporate terms whose meanings presuppose the truth of some comprehensive deterministic physical theory, can completely describe that which exists at the instant in question.

No one denies of course that it is conceivable that propositions, P say, which specify (a) relevant deterministic physical laws, (b) so-called initial conditions, at some instant, can logically imply propositions, P₂, which describe some subsequent state of affairs. But to acknowledge the conceivability of *this* is not to acknowledge the conceivability of (1), for of course the propositions P merely *incorporate* a description of what exists at the instant in question. The propositions P specify, not just the initial conditions, but also physical laws. In order to establish that (1) is conceivable, it is necessary to establish that it is conceivable that propositions, P₁, which specify or describe no more than the initial conditions can logically imply propositions, P₂, which describe some subsequent state of affairs. Or, more precisely, in order to establish that (1) is conceivable, it is necessary to establish that it is conceivable that that which exists at some instant—the initial conditions—is such that only propositions which incorporate terms whose meanings presuppose the truth of a comprehensive, deterministic physical theory can completely describe that which exists at the instant in question.

In order to clarify precisely what it is that I am proposing here, let us suppose that the world consists entirely of spherical, perfectly elastic particles, which interact by means of three kinds of forces, gravitational, 'electrical' and elastical. Suppose that the following somewhat crude

Newtonian, atomistic, comprehensive physical theory, T, is true. The postulates of T are as follows.

(i) If a material has elasticity E (with Young's modulus = Y, Poisson's ratio = σ , Y and σ being specific constants), then:

(a) $F = \frac{Y \cdot \Delta L \cdot A}{L}$, where L is the unstressed length of a rectangular rod of the material, ΔL is the increase in length corresponding to a force F, A = unstressed area of cross section of the rod, and Y = Young's modulus.

(b) $\frac{\Delta w}{w} = -\frac{\sigma \cdot \Delta L}{L}$, where w and L are the unstressed width and length of a rectangular rod of the material, Δw is the decrease in width corresponding to an increase in length ΔL , and σ = Poisson's ratio.

(ii) For any two particles with charges $\pm q_1$, $\pm q_2$

$$F = \frac{(\pm q_1) \cdot (\pm q_2)}{d^2}$$

(iii) For any two particles with masses m_1 and m_2

$$F = \frac{G \cdot m_1 \cdot m_2}{d^2}$$

(iv) $\bar{F} = m \cdot \bar{a}$

(v) If two forces $\bar{F}_1$ and $\bar{F}_2$ act on a body, then the resultant force $\bar{F} = \bar{F}_1 + \bar{F}_2$

(vi) Any particle of mass m, charge $\pm q$, unstressed radius r, elasticity E, interacts with any other particle with mass m, charge $\pm q$, unstressed radius r, elasticity E in such a way that m, $\pm q$, r, E remain constant.

(vii) The world consists entirely of particles mass m, charge $\pm q$, unstressed radius r, elasticity E.

Granted that T is true, then propositions, formulated in terms of T, which specify initial conditions, which describe what exists at any instant, will be of the following type: 'At x, y, z, t , there is a particle, mass m, charge $\pm q$, elasticity E, velocity v, which is a sphere of radius r when unstressed, but whose present shape and size is such and such'. Now suppose that that which exists at some instant is such that it can only be completely described if the descriptive terms employed, e.g. 'E', 'm', 'q', *presuppose the truth of the postulates* (i)-(vi) of T. Suppose, in other words, that that which exists at some instant is such that it can only be completely described in terms of T if the meaning of the terms 'E', 'm', 'q' are interpreted in such a way that the postulates (i)-(vi) of T are *analytic* statements. In this case a complete description of what exists at some instant, formulated in the terms of T, will presuppose the truth of postulates (i)-(vi) of T, and

will thus logically imply descriptions of subsequent states of affairs. In other words, that which exists at some instant will be such that any set of propositions which gives a complete description of this, will logically imply propositions which describe subsequent states of affairs.

It should be noted that the fact that the postulates (i)-(vi) of T are analytic does not mean that T itself is analytic, for of course (vii) is non-analytic. The entire empirical content of T is contained in the postulate (vii). (In what follows 'T' is to be understood in such a way that the postulates (i)-(vi) are *analytic* statements.)

Terms of T such as 'mass', 'elasticity', "'electrical" charge' and 'gravitational charge' are intended to be analogous to such ordinary observational terms as 'inflammability', 'solidity', 'solubility', etc., in that they are intended as *descriptive* terms, as terms that can be used to ascribe a property to an object, to describe what exists at some instant. Of course a T-term differs from any ordinary observational term in that the meaning of a T-term presupposes the truth of a precise, comprehensive theory, whereas the meaning of a term such as 'inflammability' presupposes the truth of only some extremely limited and imprecise theory. I wish to argue, however, that this difference does not justify one in maintaining that only observational terms, and not also highly theoretical terms such as those of T, can be descriptive, in the above sense.

I must now consider some objections to the thesis that (a) a physical theory such as T is conceivable, and (b) if such a physical theory is conceivable, then (1) is conceivable, or possibly true.

In the first place it may be argued that a physical theory that ostensibly postulates and describes certain unobservable, theoretical entities, cannot be interpreted as *really* postulating and describing these entities, at least in so far as the theory is considered as a theory of *physics*. The job of a physical theory is to predict observable results of experiments, and this becomes possible only when the theoretical terms of a theory have been related to, or interpreted in terms of, observational terms by means of 'correspondence rules'. Thus a physical theory is 'about', not the theoretical entities it ostensibly postulates and describes, but rather the observable results that it predicts. Hence propositions, formulated in terms of the uninterpreted theoretical terms of a physical theory, that ostensibly describe initial conditions, cannot be regarded as really describing anything at all. Only propositions, formulated in terms of observational terms can be interpreted as describing initial conditions, i.e. something that does actually exist.

There is not space for me here to give a reply to this objection. I shall assume, without discussion, that there is no valid, *general* objection to

giving a tentatively realistic interpretation to physical theories.¹ I shall assume that a physical theory such as *T*, formulated entirely in 'theoretical' terms, unrelated to any 'observational' terms by means of 'bridge' statements, or 'correspondence rules', can legitimately be interpreted as *really* postulating and describing the entities that it *ostensibly* postulates and describes. I assume, in other words, that 'theoretical' terms of physics, in addition to 'observational' terms, can be interpreted as *descriptive* terms, as terms intended to refer to entities or properties conjectured to exist.² On this view, 'bridge' statements, connecting 'observational' and 'theoretical' terms, constitute synthetic *identity* statements: they assert that certain observational entities or properties are, as a matter of contingent fact, certain theoretical entities or properties.³

Granted that there are no valid, general objections to interpreting a physical theory such as *T* realistically, it may nevertheless be objected that it is not conceivable that the particles, postulated by *T*, could have, at any given instant, the kind of properties attributed to them by *T*. At most, it is conceivable that these particles could have, at any given instant, properties attributed to them by a theory *T'*, where *T'* is similar to *T* in every respect except that at least some of the postulates (i)-(vi) are synthetic, not analytic. But propositions, formulated in terms of *T'*, that describe initial conditions, that state what exists at some instant, cannot alone logically imply descriptions of subsequent states of affairs. Hence even if it is granted that it is conceivable that a true, realistically interpreted, deterministic, comprehensive physical theory can be formulated, this does not guarantee that (1) is conceivably true.

My reply to this objection is as follows. If *T'* is interpreted realistically then it follows *analytically*, from what one means by 'realistic interpretation', that at least some of the postulates of *T'* are analytic. For in asserting that *T'* is interpreted realistically, one is asserting that *T'* postulates the existence of, and ascribes properties to, particles of unstressed radius *r*. Clearly *T'* must attribute some kind of property to the particles if they are to be differentiated from empty space. Let us suppose that *T'* attributes

¹ Of course in *particular* cases there may well be valid objections to interpreting a physical theory realistically: many valuable physical theories, if interpreted realistically, become incompatible with other accepted physical theories, or internally inconsistent.

² One main objection to the thesis that *in general* physical theories cannot be interpreted realistically is that this thesis depends on the possibility of making a *general, precise, non-arbitrary* distinction between theoretical and observational terms, and no such distinction is, it seems, possible. One main reason why it is *desirable* to interpret physical theories realistically is this: there is an important sense in which only an 'essentialist'—and hence realistic—theory can be an *explanatory* theory (see p. 17, n. 2).

³ For two recent accounts of 'bridge' statements along these lines see: K. F. Schaffner, 'Approaches to Reduction', *Philosophy of Science*, 34 (June 1967), 137-47; L. Sklar, 'Types of Inter-Theoretic Reduction', this *Journal*, 18 (1967), 109-24.

270929

elasticity E to the particles. But the assertion that a certain particle has elasticity E (with Young's modulus = Y , Poisson's ratio = σ) must imply that that particle 'obeys' the postulates (ia) and (ib). And this in turn implies that (ia) and (ib) are *analytic* statements, since the meaning of 'elasticity E ' presupposes the truth of (ia) and (ib). Furthermore the *meaning* of 'elasticity E ' must presuppose the truth of some 'conservation' postulate concerning elasticity, for if the elasticity of a particle varied in an arbitrary manner, it would be meaningless to attribute elasticity E to that particle at some instant.

To this it might be objected that in attributing the dispositional property of elasticity to the particles, one might not mean anything so precise as that the particles 'obey' (ia) and (ib). One might mean simply that when the particles collide, they rebound, momentarily altering their shape.

But what possible justification can there be for maintaining that at any instant an object can only have a somewhat *vague* dispositional property? How could one decide whether a given dispositional property is sufficiently vague to be capable of existing at an instant? Clearly the thesis that 'elasticity E ' is too precise a dispositional property to be capable of being possessed by a particle at an instant, must be rejected.

Once it is acknowledged that a particle can have the dispositional property of elasticity at an instant, what justification can there be for denying that the particles can have the dispositional properties of mass, 'electrical' charge and gravitational charge? If a particle can have, at an instant, a property which can only be referred to if (i) constitutes an *analytic* statement, why should not the particles have, at an instant, properties which can only be referred to if (ii)-(vi) are analytic statements?

In short, once it is granted that T can legitimately be interpreted realistically, there can be no more reason for denying that the particles, postulated by T , can have, at any given instant, the dispositional properties of elasticity, mass, 'electrical' charge and gravitational charge, then there can be for denying that petrol can have, at any given instant, the dispositional property of inflammability, or diamond the dispositional property of solidity.

As I have already remarked, the fact that the postulates (i)-(vi) of T are analytic does not mean that T itself is analytic, for of course (vii) is a non-analytic statement. It is true that if T is to be a falsifiable theory, then non-analytic 'bridge' statements must be added to T . But these 'bridge' statements may be regarded as falling within the province of postulate (vii), since they state, when taken together, that all the various observational entities and properties known to exist are in fact certain combinations of entities and properties postulated by T . In other words the 'bridge'

statements may be regarded as a partial, but falsifiable, rendering of the statement: The world consists entirely of particles of mass m , charge $\pm q$, unstressed radius r , elasticity E . Clearly T can be a falsifiable theory even though postulates (i)-(vi) are all analytic.

There is a sense in which T , if it were true, would provide an ultimate explanation of phenomena, an explanation whose basic premises could themselves require no further explanation. But, it may be objected, the task of theoretical physics can never be completed in this kind of way, for at any level there must always remain further legitimate theoretical problems to be solved. Hence it would be extremely bad policy for the physicist to formulate any theory analogous to T in the above respect. T , for methodological reasons, must be characterised as a non-scientific, metaphysical theory.¹

But even if T is true, and does, therefore, provide an ultimate explanation of phenomena, this could never be known for certain. Thus even if T meets with complete empirical success, it will still be desirable for physicists to attempt to refute T , to attempt to devise theories which are not refuted where T perhaps is refuted. One could not be certain, however, that such a programme would meet with success. Just as no-one can ever be certain that the 'ultimate', true physical theory has been formulated, so equally no-one can be certain that the 'ultimate', true physical theory will never be formulated. I conclude that it would not be methodologically inadvisable for the physicist to formulate a theory analogous to T , supposing such a thing were possible.

A somewhat different objection to the thesis that the truth of T implies the truth of (1) is the following. Suppose a true, comprehensive description, D , of the universe is given in terms of T at some instant. Then D cannot logically imply a description of any subsequent state of affairs, for it is always logically possible, given the truth of D , that entities should suddenly 'pop' into existence and interact with the particles described by D .

This objection can be met if T is reformulated so that it, in effect, attributes dispositional properties to space, i.e. if T is reformulated as a field theory. For if any point of empty space has the property that an entity can exist there only if it is a particle postulated by T , which moves in accordance with the 'laws' of T , then at no point in space can an entity suddenly 'pop' into existence. A comprehensive description of the universe at some instant would not logically permit entities to 'pop' into existence at subsequent instants.

¹ Popper, for example, has argued that ultimate, or what he calls 'essentialist', explanations ought to be excluded, for methodological reasons, from science; see K. Popper, *Conjectures and Refutations* (Routledge and Kegan Paul, London, 1963), pp. 104-7.

Finally it might be objected that if the argument of this section is accepted, then for any two successive events E_1 and E_2 , in any universe whatsoever, (1) will be true, since it must always be possible to formulate propositions P_1 which constitute a description of what exists at the time of occurrence of E_1 , and which logically imply propositions P_2 which state that E_2 occurs subsequently. But if (1) is *analytically* true, (1) cannot be interpreted as making an *existential* claim.

My reply to this is as follows. If the propositions P_1 (a) are to qualify as constituting no more than a description of what exists at t_1 , the time of occurrence of E_1 , and (b) are to imply logically P_2 , then, at some level, P_1 must attribute, to entities at t_1 , dispositional properties which cannot, by definition, change. In other words P_1 must incorporate terms which (a) ascribe dispositional properties to entities at t_1 , (b) are such that it follows analytically, from the meaning of these terms, that the dispositional properties in question do not change. (Derivative dispositional properties may of course change. But, at some level, the manner and circumstances in which these dispositions change must not themselves change; and this implies the existence of higher level unchanging dispositional properties.) Clearly, unchanging dispositional properties cannot be truly ascribed to objects in any universe whatsoever.

IV

Finally, what are the *implications* of rejecting Hume's contention? This question really needs another paper. I can here do little more than indicate what seem to me to be the implications of rejecting Hume's thesis.

When one first reads Hume one is inclined to conclude, considerably shocked, that if Hume's arguments concerning the notion of 'necessary connection' are correct, then:

(7) There can be no rational grounds for the belief that the orderliness, regularity, or *lawfulness* that has been found in the world so far will persist in the future.

(8) Any orderliness, regularity, or lawfulness that does persist must be utterly inexplicable, must be something for which there can be no explanation whatsoever.

Hume's arguments can seem to imply this because one assumes that:

(9) There can only be rational grounds for the belief that past lawfulness will persist in the future if it is rational to believe that that which exists at one instant determines, with logical necessity, what exists subsequently.

(10) There can only be a genuine explanation of observed, persistent regularities if that which exists at one instant determines, with logical necessity, what exists subsequently.

It has often been pointed out, however, that Hume's contention does not entail (7) and (8) above. For if Hume's contention is *true*, then it is *meaningless* to assert 'That which exists at one instant determines, with logical necessity, what exists subsequently'. In other words the stipulations (9) and (10) become *meaningless*. One must therefore revise one's notions of what can constitute necessary conditions for (a) rational grounds for the belief that past regularities will persist in the future, (b) an explanation of persistent observed regularities.

On the other hand, if Hume's contention is *false*, then the stipulations (9) and (10) are no longer meaningless. Thus the need to revise one's notions of what can constitute necessary conditions for (a) and (b) above collapses. Both (9) and (10) become legitimate assumptions. And it seems to me that if (9) and (10) *can* be accepted then they certainly *ought* to be accepted.

In short, once one rejects Hume's contention concerning necessary connections, then one must reject also those quasi-Humean theses whose acceptability depends on the *validity* of Hume's contention. It seems to me that quasi-Humean accounts of nomic universality, scientific explanation, and (a) and (b) above, all fall into this category.

University College, London

Scientific Revolution for Ever?

by WILLIAM KNEALE

I

Within the last few years a number of philosophers of science have suggested that it would be a sad ending of scientific endeavour if mankind ever came to permanent acceptance of a single theoretical framework. In their view intellectual life is at its best when each hypothesis is succeeded after a while by a better, so that science never reaches stability. Such a programme of perpetual revolution is expressed clearly in the following passage from the beginning of a recent essay by Sir Karl Popper called 'Truth, Rationality and the Growth of Scientific Knowledge':

I assert that continued growth is essential to the rational and empirical character of scientific knowledge; that if science ceases to grow it must lose that character. It is the way of its growth which makes science rational and empirical; the way, that is, in which scientists discriminate between available theories and choose the better one or (in the absence of a satisfactory theory) the way they give reasons for rejecting all the available theories, thereby suggesting some of the conditions with which a satisfactory theory should comply. You will have noticed from this formulation that it is not the accumulation of observations which I have in mind when I speak of the growth of scientific knowledge, but the repeated overthrow of scientific theories and their replacement by better or more satisfactory ones ([4], p. 125).

Everyone will agree in hoping that scientific progress does not come to an end with a false theory, as might happen if the human race were destroyed or reduced to such misery that it could no longer find leisure for research. But some may still entertain the seventeenth-century hope that scientific progress will come to an end with an all-embracing true theory. It is clear that Popper is not attracted by such a prospect, but it is not so clear whether he thinks it a real possibility. There are on the other hand philosophers of science who start from the assumption that there can be no finally satisfactory theory, because nature is infinitely complex, and argue for this reason in favour of perpetual scientific revolution. Professor David Bohm, for example, has said in a symposium of philosophers and physicists:

Received 1.x.67.

The main point of science is not so much to make a summary of the things one knows, but rather to find those laws which are essential and which therefore hold in new domains and predict new facts. . . . Generally speaking, by finding the unity behind the diversity, one will get laws which contain more than the original facts. And here the infinity of nature comes in: the whole scientific method implies that no theory is final. It is always possible there is something that one has missed. At least as a working hypothesis science assumes the infinity of nature; and this assumption fits the facts much better than any other point of view that we know ([2], p. 56).

What Bohm had in mind when making this short comment is revealed more fully by a contribution of Professor Jean Paul Vigiér to the same symposium. Writing of the attempts he and Bohm had made to reconcile continuity and discontinuity in physical theory, he said:

We would prefer to say that at all levels of Nature you have a mixture of causal and statistical laws (which come from deeper or external processes). As you progress from one level to another you get new qualitative laws. Causal laws at one level can result from averages of statistical behaviour at a deeper level, which in turn can be explained by deeper causal behaviour, and so on *ad infinitum*. If you then admit that Nature is infinitely complex and that in consequence no final state of knowledge can be reached, you see that at any stage of scientific knowledge causal and probability laws are necessary to describe the behaviour of any phenomenon, and that any phenomenon is a combination of causal and random properties inextricably woven with one another. All things in Nature then appear as a dialectical synthesis of the infinitely complex motions of matter out of which they surge and grow and into which they finally are bound to disappear ([2], p. 77).

In what follows I shall first consider how the programme of perpetual revolution in science is related to earlier accounts of scientific method, and in particular to Popper's; then I shall examine the doctrine of the infinite complexity of nature which would, if true, be the strongest justification of the programme; and finally I shall discuss Popper's dislike for stability in order to see whether his legitimate demands can be met without perpetual revolution.

II

There are two independent sources for our common stock of ideas about scientific method. The first, which may be traced back to Aristotle's remarks about induction, is the tradition of science as a body of *generalisations* about the behaviour of observable things, i.e. such generalisations as we have in elementary chemistry and biology. The second, which may be traced back to Eudoxus, is the astronomical tradition of *explanation* by means of hypotheses which preserve the phenomena, as Simplicius says in his commentary on Aristotle's *De Caelo* ([5], p. 509 ff.). In the

seventeenth century the first tradition was confused by Bacon's extremely influential use of the word 'induction' for a procedure that is supposed to lead to the discovery of causes, or generative forms, by elimination; but in our time pure, or primary, induction is once more taken to be argument for universal conclusions by the citing of cases which fall under them. In the latter half of the same century the second tradition was enlarged by Boyle, Hooke, Leibniz, and Huygens into a general theory of physical explanation by means of hypotheses about unobservables such as atoms or ether waves. But this extremely promising development of methodology failed to get the attention it deserved from philosophers of the eighteenth century, because Newton's authority, like that of Berkeley and Hume, seemed to favour a positivist programme of generalisation without explanation. It was only with the work of Whewell and Jevons in the nineteenth century that the hypothetical method came once more into honour, and even then its triumph was slow. As late as 1858 Whewell himself described the history of physics in Baconian style as the rise from less to more general laws and warned his readers against supposing that the success of atomic theory in 'expressing and calculating laws of nature' was any argument for its truth ([6], p. 19).

Already in seventeenth-century writing about the hypothetical method the old astronomical term 'phenomenon' was used freely with the extended sense of 'generalization about observables', as when, for example, Leibniz wrote to Conring in 1678 'It is the greatest commendation of an hypothesis if by its help predictions can be made even about phenomena or experiments not yet tried' ([3], vol. 1, p. 196). For obviously what is supposed to be deduced from a scientific hypothesis of the kind Leibniz had in mind is not a singular proposition about a datable occurrence but a law about the behaviour of all observables of some kind, and the deduction of the law is called a prediction only because it is made before the confirmation of the law by induction from observations. In other words, the hypothetical method is not supposed to be an alternative to primary induction, but essentially a supplement to it, and that is why I once proposed that it should be called secondary induction ([1], p. 104).

Sometimes Whewell and Jevons use the word 'induction' in a way which suggests that they think it unnecessary to make any distinction such as I have drawn by use of the adjectives 'primary' and 'secondary'. But this seems to me unfortunate, because the distinction is needed for making the interesting point about prediction to which I have just drawn attention. While we may say, quite correctly, that the propositions which Leibniz calls phenomena are no more than very well confirmed hypotheses in that wide sense in which all empirical propositions other than records of

particular observations may be called hypotheses, we must allow that they differ in an important way from the propositions for which Leibniz reserved the title of hypotheses, since they can be tested through their instances, whereas the others can be tested only less directly, through them. That is exactly why the old astronomical terminology of 'phenomena' and 'hypotheses' seems appropriate for the extended usage it received in the seventeenth century. In the astronomical sense phenomena are already generalities, though concerned with individual heavenly objects such as the sun, moon, and planets; and the reason why they, unlike hypotheses, must be preserved through all speculation is simply that they can be checked independently by repeated observations from this earth, whereas hypotheses such as that of Eudoxus about homocentric celestial spheres can never be checked independently by earth-bound mortals. Although he was extremely bold in his intellectual constructions, Eudoxus seems to have regarded speculation in astronomy as useless except in so far as it could be controlled by observation. That I take to be the point of the story that he refused to make guesses about the composition of the sun though he said he would gladly share the fate of Phaëthon if he could learn the truth by going nearer. Perhaps it was also the reason for his writing a book of descriptive astronomy under the title *Phaenomena*, from which the technical usage seems to be derived. Similarly in natural science phenomena are more fundamental than hypotheses from the epistemological point of view, though they may be not at all fundamental in the order of nature. But here hypotheses are distinguished from phenomena not merely as propositions which we cannot check directly because of our inability to travel freely in space but as propositions concerned with objects such as atoms or ether waves which are supposed to be essentially unobservable.

Since the great expansion of knowledge in the last century philosophers who are interested in scientific method have often speculated about the pattern of scientific progress, and some have written as though all the commonly recognised sciences may be expected to develop according to one set plan. It has been suggested, for example, that just as many phenomena have been explained successfully in physics by hypotheses about unobservables so generalisations of biology or psychology may presently be explained by hypotheses about theoretical entities proper to biology or psychology, as the case may be. We cannot rule out the possibility of such development, but we have no good reason to suppose that every study we distinguish from time to time for the purposes of academic organisation will continue to lead an independent life. On the contrary, we can see already that chemistry is best conceived as a branch of physics, and we may reasonably suppose that the advance of science will lead to the removal of

other traditional boundary fences. As for the contrast in physics between observables and unobservables, this, like the connected distinction between primary and secondary qualities, is grounded in the nature of sense perception and needs therefore to be made only once for all studies. Since these points are well known to Popper, it must be assumed that when he talks of the growth of scientific knowledge he is thinking of progress in science as a whole, rather than of little improvements in a number of separate sciences; and he provides clear evidence for this interpretation in the paper from which I have quoted when he writes that the new theories which interest him are concerned with 'connexions or relations between hitherto unconnected things'. Similarly Bohm and Vigier, though they are physicists by profession, conceive their science in the old way as the all-embracing natural philosophy. So the scientific revolutions which we are going to discuss must be understood to be comprehensive totalitarian affairs.

From time to time in the course of the development I have sketched philosophers have maintained the following three theses which seem to me mistaken: (1) that induction is a method of discovery, capable of extracting new general truths from systematic records of experience; (2) that the progress of inductive science consists precisely in the accumulation of such truths; (3) that inductive inference from genuine observations can attain, if not certainty, then at least high probability in the sense of the theory of chances. The first two, which are to be found in Bacon's work, have long since been discredited by champions of the hypothetical method. For it is obvious that there is no algorithm for the devising of good explanatory hypotheses about unobservables, and even the true generalisations about observables which are commonly called phenomena cannot be discovered by rules of thumb, since we notice only what interests us. When there is already a beginning of theory, most of the observations made by scientists are for the purpose of testing the phenomenal consequences of various thinkable developments. And the progress which especially interests practising scientists is not the lengthening of their inventory of phenomena, but precisely that theoretical development from which the lengthening results. The third of the theses I have listed is not a contention of Bacon. On the contrary, such talk of probability in the sense of the theory of chances belongs rather to the tradition of hypothetical reasoning in natural science, and unlike the first two, it has some defenders among modern philosophers. In my view it is mistaken because talk of chances is meaningful only when it is reasonable to speak of alternative possibilities within a well-established theoretical framework. While it is right to assert against Bacon and Mill that ampliative induction yields probability rather than

certainty, it is not right to suppose that there can be chances for or against propositions about the frame of nature such as we try to establish by induction.

Popper has often declared himself an anti-inductivist, and part of what he means by this is certainly that he rejects the three theses I have just noticed. But with a degree of zeal which I find very puzzling he has pressed his criticism in each case to the point of paradox, saying (1) that there is no such thing as a good inductive argument, even for support of a conclusion suggested by theory, (2) that there can be no accumulation of positive, as distinct from negative, results in the progress of science, and (3) that it is a mistake to talk at all about the probability of scientific results. Clearly these denials hang together. Having committed himself to one, he can scarcely refrain from the others. And it is interesting to notice that together they might conceivably suggest a nihilistic doctrine of perpetual revolution in science. For if it is a mistake to suppose with the inductivists that there can be solid gains from scientific reformism, we have nothing to lose by scientific revolution but our chains of error, and we may as well enjoy what excitement we can get from destroying the hateful things. For this purpose it will be desirable, of course, to work on intellectual chains of the best quality, that is to say on theories which are strong in explanatory power and not easy to refute. As Popper declares in one place, we shall require of a good theory 'that it is not refuted too soon' ([4], p. 247), though all must yield in the end to our destructive *élan*.

In general, however, Popper does not present his new programme as nihilistic. On the contrary, like Bohm, he seems anxious to assure us that it can bring all the blessings of liberalism. In particular what they say or assume about inductive argument, accumulation of phenomena, and the probability of hypotheses is nearer to ordinary scientific views than to the earlier, more extreme, contentions of Popper.

First, Bohm and Popper follow Leibniz in saying that it is a mark of a good theory to allow prediction of hitherto unsuspected phenomena. In a passage which I quoted at the beginning of this paper Bohm talks of predicting new facts, and in Popper's essay there is a similar passage where he says that we require of good theory that it be successful in predicting some new effects ([4], p. 247). Here the words 'fact' and 'effect' are used, as often in science, with the sense of 'phenomenon', and clearly both authors are right in what they assert. But how can we ever say that a theory is successful in any of its predictions unless we assume that these may be confirmed inductively? Perhaps Popper has overlooked this difficulty because the phrase 'prediction of an effect' could be used in other contexts for a statement in the future tense that was verifiable conclusively by a

single observation. Or it may be that he is aware of the difficulty of reconciling what he has written with anti-inductivism but thinks it can be defended by saying that success may be claimed for a theoretical prediction on the basis of a single observation because prediction of a phenomenon, or effect in the technical sense, may be understood as prediction of each of the cases it covers. It should be noticed, however, that if a scientist has predicted a certain phenomenon on theoretical grounds we do not count a single favourable observation as a success unless we also think that the conditions of the observation are such as to justify us in asserting the universal proposition which was in fact derived from the theory. We should not say, for example, that a single observation of stars during an eclipse of the sun in 1919 verified a theoretical prediction of Einstein if we did not think that it virtually established a hitherto unsuspected law about the apparent displacement of distant bodies seen or photographed over the edge of a massive body. In short, the only theoretical interest of the observed happening is the interest it gets from being a case of the predicted law which we now believe ourselves entitled to assert.

Secondly, Bohm and Popper assume that phenomena which are well confirmed by primary induction remain as scientific results even after the refutation of the hypotheses by which they were originally suggested. This is implicit in their claim that scientific progress leads to hypotheses of greater explanatory power. For we cannot be sure that a new hypothesis has greater explanatory power than its predecessor unless it entails all the genuine phenomena entailed by the predecessor together with at least one other. Naturally we do not wish a new hypothesis to entail the false predictions of its predecessor, and so it is not enough to say simply that scientific progress leads to hypotheses of wider scope. If there is indeed progress in spite of the ultimate refutation of all theories, we can have no assurance of this except from an increase in the number of genuine phenomena explained.

Thirdly, although Popper still refuses to talk of the probability of theories, he now suggests that scientific progress involves transition to hypotheses of greater verisimilitude. From his definition of his new concept of verisimilitude it is clear that he is not concerned with likelihood, *vraisemblance*, or *Wahrscheinlichkeit*, that is to say with appearance of truth, but with approximation to truth, which he supposes to be the original sense of the similar Greek word *eikos*. I do not think he is right in saying, as he does, that 'confusion between verisimilitude [in his sense] and probability goes back almost to the beginning of Western philosophy', and that this is the source of 'the feeling that it is absurd to deny that science aims at probability' ([4], p. 237). In my view the real source of Popper's

worry is his obsession with the technical use of the word 'probability' for a proportion of chances, a use which appears to have been introduced by De Moivre as late as 1718. But I think it is very interesting that Popper now recognises the antiquity of the thesis he attacks and tries to explain it away as a misunderstanding of a thesis he accepts. Bohm does not discuss these questions, but I have the impression that he would be happy to work with the notion of approximation to truth for which Popper proposes to reserve the word 'verisimilitude'. Whether either is entitled to assume that an unending sequence of scientific revolutions can provide infinitely many successive approximations to truth is a question to be considered in the next section.

III

One of the doctrines which Marxists claim to inherit from Spinoza is that of the infinity of nature. Ordinarily this is taken to mean that any cosmology which denies the infinite extension of the universe must be wrong. But it can be interpreted in other ways, with some authority from Spinoza himself. So perhaps it is not surprising to find Bohm and Vigier asserting as a postulate of scientific thought that nature is infinitely complex and therefore not to be comprehended in any finally satisfactory theory. We must consider presently how they apply their principle in detail, but before we come to that it will be useful to distinguish their notion of a sequence of scientific theories from two other possibilities with which it might be confused.

Whenever an explanation is given for a fact or group of facts, we may find the explanation puzzling and look for yet another explanation at a deeper level, and when we get that we may look for a third, and so on. If, however, each new explanation has no consequences other than those which follow from its predecessor, it is really no more than a logical or mathematical transformation of the first. That is to say, all the explanations in the series are equivalent, and there is nothing to choose between them, except that some are simpler to grasp than others. This difference is not negligible, as everyone knows who has learnt something of the technique of simplifying algebraical expressions, but it is implausible to suppose that there might be an infinite regress of expressions each simpler than its predecessors according to the same standard of simplicity, and in any case what we have here is not what Bohm and Vigier expect in natural science. For in their programme each new theory is required to have consequences that could not be derived from its predecessors, and it is therefore impossible to think of it as no more than a transformation of its predecessors.

Something nearer to their requirement is a regress of historical explanations like that given in the nursery rhyme about the loss of kingdom all for the want of a horseshoe nail. For here every event mentioned as a cause must be supposed to have effects other than that it is required to explain; and so far as I can see there is no reason why a temporal regress of causes should not be infinite, though we can never trace more than a finite part of it. But again we have no exact analogy here to what Bohm and Vigier want. For according to their programme each unifying hypothesis put forward by a scientist must be refuted in due course unless the progress of science is stopped by some misfortune. Certainly the phenomena explained by each hypothesis are to be explained also by its successor, together with something new, but for each hypothesis there is to be at least one false phenomenal consequence through which it can be refuted, and their view of the nature of scientific progress requires that this too should be a different consequence from any entailed by earlier hypotheses. So the various hypotheses of their sequence are not to be supposed nested in the sense that each is entailed by its successors. To assume that would be to adopt for the realm of hypotheses, as well as for the realm of phenomena, what Popper calls the inductivist account of scientific growth by accumulation.

A simplified model of a Bohm-Vigier sequence of hypotheses may be constructed as follows. Let us suppose that $P_1, P_2, P_3, \dots$ are phenomena in the relevant sense and that $F_1, F_2, F_3, \dots$ are false phenomenal propositions, i.e. false generalities about the behaviour of observable things. Each set may be assumed infinite and capable of being well-ordered in some way. Then if C_n is taken to be the conjunction $P_1 \& P_2 \& P_3 \& \dots \& P_n \& F_n$ for every positive integer n , the sequence $C_1, C_2, C_3, \dots$ will be found to have most of the required properties. Each member entails all the phenomena entailed by its predecessors, together with a new truth and a new falsehood. And given the assumption, which I am prepared to grant, that the sequence $P_1, P_2, P_3, \dots$ and $F_1, F_2, F_3, \dots$ are both infinite, there is no difficulty in taking the sequence $C_1, C_2, C_3, \dots$ to be infinite also. It is true that Bohm and Vigier do not say each new hypothesis entertained by scientists should entail just one new phenomenon and one new false phenomenal proposition, but they do insist it should have at least these two new consequences. For a slightly more realistic model we may, if we like, suppose that scientific progress involves great leaps down the sequence $C_1, C_2, C_3, \dots$, e.g. from $C_{1,000,000}$ to $C_{2,000,000}$; and by introducing a few more complications we can even allow for leaps of infinite length without removing the possibility of unending progress. Furthermore all these models meet Popper's requirement of increasing verisimilitude, since the balance of truth content over falsity content grows

steadily as we progress through the sequence of conjunctions. I have not tried to work out a more complicated form in which Popper's requirement would be satisfied even though each new hypothesis entailed an infinity of false phenomenal propositions, but I dare say it could be done with a little ingenuity.

So close indeed is the fit that I suspect Bohm, Vigier, and Popper may sometimes have allowed their thought to be guided by such a scheme as I have suggested. But none of them would wish on reflection to allow that a scientific hypothesis can be identified with the conjunction of its phenomenal consequences, since a mere conjunction has no explanatory power. On the contrary, they all maintain, quite rightly, that genuine explanation always involves a new unifying idea, and sometimes they seem to be committed also to the thesis that a good overall theory must be an organic unity in the logical sense, i.e. not resolvable into a conjunction of items, whether phenomenal or non-phenomenal, that can be examined separately. For if a promising theory could be resolved in this way, when its time came for refutation we might be able to trace the fault down to some separable part or parts and so save the rest for incorporation in a new conjunction, which might make for increase of verisimilitude but would be a reformist rather than a revolutionary proceeding. In fact the difficulties of the doctrine of the infinite complexity of nature appear only when we leave behind the oversimplified model I presented a few minutes ago and try to combine the notion of infinite complexity with that of theoretical unification.

The main trouble may be put shortly as follows. There seems to be no good sense in which nature can be said to be so complex as to require perpetual revolution in science. For if by 'the infinite complexity of nature' is meant only the infinite multiplicity of the particulars it contains, that is no bar to final success in theory making, since theories are not concerned with particulars as such. So too, if what is meant is only the infinite variety of natural phenomena, since that may be comprehended in a unitary theory, just as infinitely many different mechanical phenomena are covered by Newton's law of gravitation. If, however, what is meant is precisely what Bohm and Vigier state as a consequence of the principle, namely that there cannot be a finally satisfactory theory, there is no point in trying to approach such a theory, either by scientific reforms or by scientific revolution. For if there is no truth, there cannot be any approximation to truth, and the whole enterprise of theory making is futile. Nor does it help to say that there is indeed a true explanatory theory in some Platonic heaven but that it is infinitely complex and so not to be comprehended by men. For if there can be an infinitely complex proposition, it will certainly not be a single explanatory theory in any ordinary sense of

that phrase, but at the best an infinite conjunction of explanatory theories. Perhaps we can produce successive approximations to such a conjunction, if there is nothing else to work for, but in that case our best hope of success will be by steady accumulation of the separate items, rather than by perpetual revolution.

Since we are not directly concerned with the way in which Bohm and Vigier apply their principle, I have tried so far to discuss it in abstraction; but it will be useful to notice briefly what kinds of scientific revolution they foresee, since this will help to explain how they interpret the principle. Their view is that science may be expected to pass through alternating phases of quasi-determinism and quasi-indeterminism. First scientists try to explain all phenomena within a causal scheme. Then, when they seem to be in sight of success, they come on phenomena which convince them that their original project is hopeless, and so they switch to statistical theory but try to retain most of their first theory in the guise of an approximation to causal explanation within the new framework. At the third level new phenomena discredit the universal statistical theory of the second level and there is a return to causal theorising, but most of the theory of the second level is retained in the guise of an approximation to statistical explanation within a causal framework. At present Bohm and Vigier think they are entering the third phase, but they believe also that the sequence of possible alternations is infinite and that the business of a scientist is to carry out in such detail as he can the programme laid down in advance by their dialectical philosophy. If they are right, there is no end to the scientific journey, but we can nevertheless indicate already the direction it should take. Thus, provided there are no serious disasters for the human race during the next few million years, it may be expected that someday research scientists will be working at the 349,751st level (which, being odd-numbered, will be quasi-deterministic) while science teachers and writers of text books are still dealing with all the 349,750 higher levels as approximations to approximations to approximations etc. I find this implausible. What Bohm and Vigier say about the present situation in theoretical physics is attractive, and may be true, for all I know. But extrapolation *ad infinitum* from two revolutions, of which one has not yet got beyond the manifesto stage, does not carry conviction to an inductivist like myself, and I doubt whether it even makes sense to one who works with ordinary logic as opposed to dialectics. For although the special form of their infinite regress of explanations has the apparent merit of making it seem possible for scientific research to have a direction without having a goal, it does not remove the fundamental difficulty of the principle of infinite complexity, namely that, if the principle excludes all hope of final

theory, it does so by excluding all determinate structure such as scientists might wish to present in a final theory. On the contrary, it puts the difficulty in the bright light of paradox, since it implies that nature is indeterminate even with respect to the contradictory opposition of determinism and indeterminism.

That the principle has nevertheless won the support of two able physicists may be the result of two confusions. First there is undoubtedly a confusion in their thought about the notion of a final theory. In a passage which I have already quoted Bohm writes: 'The whole scientific method implies that no theory is final. It is always possible that there is something we have missed.' Obviously the second statement is intended to support the first. But it only goes to show that no theory is *known to be final* in that sense of 'final' which is relevant to the discussion. From recognition of the fact that even the most successful of scientific theories may yet be refuted we may reasonably pass to a call for unceasing revolutionary vigilance in science but not to a programme of scientific revolutions for ever. Secondly, there may, I suspect, be a confusion also about the notion of infinite complexity. There is a quite good sense in which any continuum may be said to be infinitely complex, since it contains parts within parts *ad infinitum*, and this is the sense in which a reader is most likely to interpret the principle of the infinite complexity of nature when he first hears of it. Later on he will find Bohm and Vigier using the principle as though they took it to imply that nature has infinitely many distinguishable levels of fine structure. But the first does not entail the second.

IV

If my argument is sound, the programme of perpetual revolution in science cannot be justified by appeal to the principle of the infinite complexity of nature. So far as I can see, a scientist may someday formulate a comprehensive theory which is wholly true and therefore incapable of refutation. I do not wish to say that this will happen. The human race may be too stupid for the task, or it may be overtaken by some natural disaster while it is still far from its scientific goal. Still less do I wish to say that, if we produce the theory we want, we can also be sure that we have done so. For in order to attain such certainty we should have to know everything about nature, including the fact that we knew everything about nature; and I believe that to be impossible in principle. But the ideal of comprehensive true theory remains unaffected by these reservations, and I think it is essential to the scientific enterprise. During the present century research has been highly organised as a professional activity, and with this development there has naturally come a tendency to assume that the work

will continue indefinitely. But that does not mean that scientists have despaired of final truth and compromised for perpetual revolution in scientific theory. Most of them are very happy if they can claim a few phenomena as their own by right of confirmation, and they believe, quite reasonably, that there are plenty more to be found in the mine from which they have got their nuggets.

It is not clear to me whether or not Popper is prepared to agree that a scientist may someday formulate a comprehensive theory which is wholly true. If he thinks this impossible, it is rather surprising that he has not said so and given his reasons for the interesting thesis. On the other hand, if he thinks it possible, his cheerfulness about the future of science is difficult to understand. For according to the assertion I quoted at the beginning of this paper success in the enterprise would destroy the rational and empirical character of science, and from this there seems to follow the gloomy consequence that, if the enterprise is feasible, nothing can save us from scientific death except our own incompetence as scientists. For my own part I am not convinced that success in theory making would be as dangerous as he fears, and I will therefore conclude my paper by examining his reasons for saying what he does.

Apparently Popper believes that if we succeeded in formulating a theory whose predictions were uniformly confirmed we should cease after a time to treat it with the caution characteristic of scientists who are still actively engaged in discriminating between available theories, and that this change of attitude would have two serious consequences. By accepting the theory as established beyond all conceivable refutation we should give it the status of a truth known *a priori*, and so our science would cease to be empirical. But because the change would be a triumph of animal faith over human criticism, our science would cease at the same time to be rational. Are these fears well grounded?

When physical theory is conceived by itself as an exercise in mathematics, that is without relation to phenomena, it seems quite natural that purely theoretical terms should be taken as defined by their theoretical role and that purely theoretical statements should be regarded as certifiable *a priori*. What is an electron, we may ask, but a thing which does whatever the postulates of the theory require an electron to do? But this situation is certainly not what Popper has in mind when he fears that science may lose its empirical character if we are ever completely successful in the scientific enterprise of theory making. What interests him is physical theory conceived in the original way as an hypothesis for the explanation and prediction of phenomena, and what dismays him must therefore be the thought that through its uniform success the theory might come to be so

incapsulated in the meanings of the terms we use for the stating of phenomena that it would no longer be possible to find language for saying anything in refutation of the theory. Is this in fact a serious danger?

From the beginning beliefs about the frame of nature have been incorporated into human linguistic usage. Consider, for example, words such as 'calf', 'foal', 'puppy', 'kitten' which are used in all languages for reference to the young of animal kinds that specially interest the speakers of the languages. Obviously the currency of these words depends on the assumption that large animals grow from small animals of the same kind, which are themselves the offspring of large animals of that kind. If *per impossibile* this assumption, which is now expressed by the technical term 'reproduction', were abandoned, we should have to keep apart the notions of cat-offspring, small-cat, and young-cat which are now happily conflated in the word 'kitten'. Similarly the common use of the word 'fall' depends on the assumption that unsupported bodies tend towards the earth. If any one said that a stone had fallen from his hand into the sky, we should find it extremely difficult to understand what, if anything, he meant by this remark, though we might perhaps succeed in making something of it in a setting of science fiction. There are indeed few, if any, words which live down to the philosophical standards of Hume and Wittgenstein, that is, are suitable for description of nature without any assumptions concerning its framework of law. In their most common application even colour words and other terms for sensible qualities presuppose regularities in the way material things affect normal human observers. It is not surprising therefore that laws of nature which have been discovered by the deliberate investigations of scientists during the last three centuries may also be made the basis of linguistic practices, if they are concerned with matters of sufficiently wide interest. When we speak of switching on the light, we use the old word 'switch' in a new way which is fully intelligible only to a person familiar with some of the phenomena that are explained by the hypothesis of electric currents. Through such terms as 'electricity' and 'atom' many new phrases of work-a-day language now have links with physical theory. Our newspapers tell us, for example, how much of the electric power used in the country comes from nuclear power stations, and politicians who have only a vague smattering of physical theory work for or against the signing of a treaty to prevent the proliferation of nuclear armaments.

For these reasons I do not doubt that words may acquire new links of meaning without change of their applications, nor yet that a formula which once expressed an empirical discovery may come later to express an *a priori* dogma. And I find it at least conceivable that a scientist may become

so accustomed to working with a successful explanatory theory that he sometimes reads the requirements of the theory into the meanings of the terms he uses to state the phenomena explained by the theory. When listening to the talk of chemists, I have sometimes had the impression that this has happened already to their professional jargon, composed, as it is, of ordinary words such as 'carbon', technical terms like 'tetrachloride', and formulae specially designed to assist in the elaboration of theory. But so long as men continue to learn their mother tongues as they do at present I do not believe it possible that they should lose all power of appreciating the status of their scientific theories as successful conjectures. For a natural language is not built into the structure of a human being before he begins to lead an independent life, nor is it acquired quickly as a fully articulated whole, like the programme of a computer. It is learnt piecemeal in use over many years and in such a way that terminology for expressing phenomena inevitably comes before terminology for expressing explanatory hypotheses. This implies, however, that the former must be basically independent of the latter, even though for some people in some parts of their lives it may be contaminated by the latter; and so there cannot be a society in which it is impossible to discuss the merits of a theoretical scheme, unless the present methods of language transmission are replaced by a brain washing technique such that a child cannot learn to say 'Pussies like milk' without learning also at the same time the principles of zoology, psychology, and physiology, together with the foundations of chemistry and physics on which these rest.

It is unlikely that Popper was thinking of such a brave new world when he expressed a fear of the results of scientific success. More probably he had in mind a sad old world differing from those we have known only by the presence of scientists in the places usually occupied by witch-doctors and inquisitors, that is, a world in which it is thought perverse to question the pronouncements of highly respected authorities. If this is indeed what he fears, we can agree that production of a completely successful scientific theory might lead to a decline of free critical inquiry. But it seems to me a great exaggeration to say that it must inevitably lead to a non-rational, non-empirical view of the world. After success it would still be necessary to pass on the achievements of science by teaching, and more than ever would depend on how this was done. Perhaps at such a time the best safeguard against complacency would be a philosophical study of the history of science. The danger, as I see it, would be that people would fail to appreciate the nature of the scientific achievement, as they often fail in our time to appreciate the nature of the perceptual achievement of mankind; and if I am right, there would be even more

need of philosophy then than there is now. But I hasten to say in conclusion that I do not think the danger is imminent.

REFERENCES

1. William Kneale, *Probability and Induction* (Oxford, 1949).
2. S. Körner, ed., *Observation and Interpretation* (London, 1957).
3. G. W. Leibniz, *Philosophische Schriften*, ed. Gerhardt.
4. K. R. Popper, *Conjectures and Refutations* (London, 1963).
5. Simplicius, *In Aristotelis Quattuor Libros de Caelo Commentaria*, ed. Heiberg.
6. William Whewell, *Novum Organum Renovatum*, 3rd edn. (London, 1958).

Definition and Discovery (II)*

by FRED WILSON

III

As a science grows, it builds up pieces of imperfect knowledge toward the desired end of perfect knowledge.¹ As each piece is added, there is definition of terms, and redefinition. The details of this process are shaped by various factors. A modification of the previous example will serve to illustrate some of these. We consider an example from the behavioural sciences; schematising will bring out the features in which we are interested.

Consider the various pieces of behaviour which are commonly grouped under the term 'angry'. Suppose these pieces of behaviour are H_1 , H_2 , and H_3 .² Since only some pieces of anger-behaviour need be present as a sufficient condition for calling a person angry, the following definition of ' x is angry' = ' Hx ' is appropriate:

$$(3.1) \quad 'Hx' \text{ is short for } 'H_1x \vee H_2x \vee H_3x'$$

Now let us suppose that the scientist is looking for a property, call it ' Gx ', the exemplification of which is a necessary and sufficient (causal) condition for someone's being angry; i.e. let us suppose he is looking for a property G such that

$$(3.2) \quad (x)[Gx \equiv Hx]$$

Assume the scientist already knows some of the necessary (causal) conditions for a person's being angry. Say he knows that

$$(3.3) \quad (x)[Hx \supset F_1x]$$

This means he knows, expanding, via (3.1), that

$$(3.3a) \quad (x)[(H_1x \vee H_2x \vee H_3x) \supset F_1x]$$

and consequently that each H_i is a sufficient (causal) condition for F_1 ; i.e., that

$$(3.3b) \quad (x)[H_1x \supset F_1x]$$

etc. Also assume that he knows

$$(3.4) \quad (x)[Hx \supset F_2x]$$

* Part I appeared in the previous number.

¹ Concerning the ideal, see G. Bergmann, *Philosophy of Science* (Madison, Wisc., 1958). See also M. Brodbeck, 'Explanation, Prediction, and "Imperfect" Knowledge', loc. cit.

² These may themselves be defined, of course. E.g. ' H_1x ' might be short for 'if x is in the appropriate set, and if x speaks English, and if x is asked "are-you-angry?", then x utters "yes"'.

Then F_2 is a necessary (causal) condition for H , and each H_i is a sufficient (causal) condition for F_2 . From (3.3) and (3.4) it follows that F_1 and F_2 are jointly necessary conditions for anger, i.e. that

$$(3.5) \quad (x)[Hx \supset (F_1x \cdot F_2x)]$$

Finally, let us assume that F_1 and F_2 appear to be not only necessary but also jointly sufficient (causal) conditions for anger, save in certain anomalous and more or less esoteric cases. That means

$$(3.6) \quad (x)[(F_1x \cdot F_2x) \equiv Hx]$$

would be true, save for those anomalous cases which show that

$$(3.7) \quad (x)[(F_1x \cdot F_2x) \supset Hx]$$

is false. The falsity of (3.7) shows that ' $F_1x \cdot F_2x$ ' cannot be the desired ' Gx ' such that (3.2) is true. But since the instances which prove (3.7) false are esoteric, it is reasonable to assume that (3.6) approximates to the desired (3.2).

This schema is, I believe, a not unreasonable representation of current knowledge about anger.

The scientist desires to discover a law of the sort (3.2). To say the same thing differently, he desires to discover a definition for a term ' Gx ' such that (3.2) is true. This desire is a regulative principle of his activities. The activities directed toward the end of defining this ' Gx ' are, moreover, guided by the other knowledge the scientist has. He knows already in (3.3) and (3.4) some of the necessary (causal) conditions of anger. He knows that F_1 and F_2 are not jointly sufficient for anger. But he also knows, since the anomalous cases are more or less esoteric, that the desired definition of ' Gx ' will probably involve ' $F_1x \cdot F_2x$ ' as an important factor. The behaviour directed toward achieving the scientist's end is, therefore, no random groping: He (the scientist) must discover in what respects the anomalous cases differ from the non-anomalous cases.

Applying the Method of Difference to certain of the anomalous cases might reveal the absence of the property F_3 which is noted to be present in certain of the non-anomalous cases. The Method of Agreement might then be applied to the various non-anomalous cases to satisfy the scientist that F_3 is a necessary (causal) condition for anger. This would lead him to add F_3 to the list of 'indicators' for anger, and to hypothesise that

$$(3.8) \quad (x)[(F_1x \cdot F_2x \cdot F_3x) \equiv Hx]$$

is the desired law (3.2), i.e. that ' $F_1x \cdot F_2x \cdot F_3x$ ' is the desired definition of ' Gx '. To test the hypothesis one would then turn to the remaining anomalous cases.

Suppose these cases show (3.8) to be false. The scientist would continue his search for yet another factor to add to the definiens of ' Gx '.

He might now notice that a factor F_4 not previously taken into account. Examining the facts, he might discover that

$$(3.9) \quad (x)[(F_1x \cdot F_2x \cdot F_3x \cdot F_4x) \equiv (H_1x \vee H_2x)]$$

is true, whereas

$$(3.10) \quad (x)[H_3x \supset F_4x]$$

is false.

At this point the scientist has a choice. He can continue his search for a definition of ' Gx ' such that (3.2) is true. Or else he can revise his definition of ' x is angry' to be

$$(3.11) \quad 'H^*x' \text{ is short for } 'H_1x \vee H_2x'$$

and define

$$(3.12) \quad 'Gx' \text{ is short for } 'F_1x \cdot F_2x \cdot F_3x \cdot F_4x'$$

leading to the restatement of (3.9) as

$$(3.13) \quad (x)[Gx \equiv H^*x]$$

On the first alternative, (3.13) does not suffice to satisfy his original desire; his search is not concluded. On the second alternative, the original desire is given up and (3.13) ends his search.¹

Which alternative he chooses will be affected by several things.

(1) There is theoretical significance. It may be that H_3 is relatively unimportant from the viewpoint of theoretical development, whereas it turns out that (3.13) has a high degree of theoretical significance. Since the scientist desires not only laws and significant concepts, but also the arrangement of these into theoretical structures, in the indicated circumstances he would opt for the second alternative. On the other hand, if the scientist had reason to believe that H_3 was important from the viewpoint of theory construction, he would opt for the first alternative.

(2) Besides theoretical significance, there are other considerations that are even more important in the case of newly emerging sciences. These have to do with the criterion on the basis of which the scientist associated the 'anger criteria', H_1 , H_2 , and H_3 into the definition of ' x is angry'.

(2a) In its customary and daily usage, 'angry' is a word whose meaning is vague in the sense that there are a great many borderline cases, etc., where it is not clear whether 'angry' is applicable or not.² When one comes

¹ One might go on (*à la* Section II of this paper) and define

$$'Kx' \text{ is short for } 'Gx \cdot H^*x'$$

which would yield, via (3.13),

$$(x)[Kx \equiv Gx]$$

and

$$(x)[Kx \equiv H^*x]$$

Whether one would make this further move would depend on considerations of the sort indicated in Section II.

² The existence of such borderline cases in no way destroys the utility of the concept so far as everyday contexts are concerned.

upon such cases, all one can do is stipulate so as to remove the unclarity. Thus, it might be perfectly clear that H_1 and H_2 are cases covered by the concept of being angry, whereas H_3 covers only borderline cases. In such circumstances the scientist might well have arbitrarily stipulated that H_3 is to be included among the criteria for being angry. If so, the difference between the definitions (3.1) and (3.11) of ' x is angry' would involve nothing more than the arbitrary stipulation of the scientist himself. That would mean, other things being equal, the desire to find the necessary and sufficient (causal) conditions for a person's being angry would be as perfectly satisfied by the law (3.13) involving H^* , as by a law involving H .

(2b) Even in their customary and daily usage words have the definitions they do because people have found that thus defined the words are useful tools. But words are used for many ends besides the stating of laws (of nature). It did not take Oxford to reveal this to the world. It may be that H_3 is included among the criteria for being angry for some reason other than significance, in the sense of occurring in laws. If H_3 is included for these other reasons, then, since the scientist is, *qua* scientist, interested in significant concepts only, it will not affect his scientific purposes if he settles for (3.13) rather than (3.2), thereby redefining ' x is angry' by (3.11).

(2c) Even if significance is not the only criterion for acceptable definitions of words in their daily and customary usage, it nonetheless is *one* of the criteria. However, the beliefs on which these judgments of significance, and, therefore, of acceptability of definitions (usages) are the pre-scientific and intuitive beliefs of the 'common man'. Such beliefs, we know full well, can be quite, even disastrously, wrong. For the beliefs of the 'common man' are both pre-scientific and prejudiced. In such circumstances, the scientist may well judge that the beliefs which led him to associate H_3 with H_1 and H_2 in defining 'angry' were nothing more than pre-scientific prejudice as to what correlated with what, that is, as to what interacts with what. If this were so, the scientist would naturally settle for (3.13).

These three things, then, namely, the facts that words of daily usage are incompletely defined, that their criteria of applicability are often assigned for reasons other than significance, and that, even when the criteria are assigned on the basis of significance, the judgments of significance are often based on pre-scientific prejudice—these three things, together with theoretical significance, will determine whether the scientist will settle for the definition (3.12) and the law (3.13), or will hold on to the definition (3.1) and continue the search for a law (3.2).

The schema above involves the dropping of criteria from the definition of 'angry'. Exactly analogous remarks hold, of course, for the addition of criteria.

As illustrations, let the following suffice. Pre-scientifically, intelligence correlates with certain behavioural patterns. As psychologists have sought to articulate and refine these pre-scientific concepts and this pre-scientific knowledge, many of the quotidian associations of 'intelligence' have been dropped. The term has been replaced by the more significant (as has been discovered) term 'IQ'.¹ In this case, certain criteria of customary usage are dropped. To this, one can contrast Freud's use of 'sexual'. For reasons of theoretical significance, Freud saw fit to redefine this term. Not only did he call behaviour sexual that was ordinarily called sexual but also many other types of behaviour. The theoretical point was that there are certain needs and corresponding goals such that (a) these needs arise and disappear at various stages in the individual's life in accordance with a biologically determined curve; (b) these needs are so interrelated that the goal which satisfies one may at times satisfy another (substitution); and (c) these needs can all with relative ease be redirected toward other goals (sublimation). Needs and goals satisfying these conditions Freud calls 'sexual'. Sexual needs are therefore to be contrasted with such needs as sleep and hunger. These do not mature and disappear; the goal of one cannot satisfy the other; they cannot be redirected. However, for reasons of pre-scientific, mainly religious, prejudice, Freud's extension of the meaning of our ordinary 'sexual' remains controversial.²

However, the best illustration occurs elsewhere in Freud, at a place where Freud gives an explicit description of the process I have indicated schematically. This occurs in his book *Jokes and their Relation to the Unconscious*.³

In this book Freud is attempting to define what it is for something to be a 'joke'. Something is classified as a joke by virtue of the psychological attitude it evokes in those who hear it. Freud, in attempting to define what it is to be a joke, is searching for those characteristics of *the thing* by virtue of which it evokes the appropriate psychological attitude. He is, in other words, looking for *laws*, for the definition of a concept significantly connected to the psychological attitude evoked.

¹ For a discussion of some of the issues surrounding the use of IQ tests, see H. J. Eysenck, *Uses and Abuses of Psychology* (Baltimore, 1962), esp. ch. I.

² This issue, and others, are discussed in G. Bergmann, 'Psychoanalysis and Experimental Psychology', *Mind*, 52 (1943), 122-40; reprinted in M. Marx (ed.), *Psychological Theory* (New York, 1951).

One might also note how non-scientific considerations entered at other points in determining Freud's use of terms. Freud (and Charcot) redefined 'illness' so as to include 'mental illness', not simply for reasons of significance (though that was part of it) but also for humanitarian reasons: the result of the persuasive definition was more humane treatment of cases of hysteria. For an excellent discussion of the role of non-scientific considerations, see T. Szasz, *Law, Liberty and Psychiatry* (New York, 1963). (Unfortunately, Szasz mars the case he is making by tying it up with a Rylean, voluntaristic, metaphysics.)

³ Trans. J. Strachey (New York, 1960).

We should thus find no difficulty in indicating the aims of any new attempt to throw light on jokes. To be able to count on success, we should have either to approach the work from new angles or to endeavour to penetrate further by increased attention and deeper interest. We can resolve that we will at least not fail in this last respect. It is striking with what a small number of instances of jokes recognized as such the authorities are satisfied for the purposes of their enquiries, and how each of them takes the same ones over from his predecessors. We must not shirk the duty of analyzing the same instances that have already served the classical authorities on jokes. But it is our intention to turn besides to fresh material so as to obtain a broader foundation for our conclusions. It is natural then that we should choose as the subjects of our investigation examples of jokes by which we ourselves have been most struck in the course of our lives and which have made us laugh the most.¹

Later on Freud indicates some of the things which guide him in his search for the relevant laws.

This is an opportunity for making a not unimportant admission. We are engaged in investigating the technique of jokes as shown in examples; and we should therefore be certain that the examples we have chosen are really genuine jokes. It is the case, however, that in a number of instances we are in doubt whether the particular example ought to be called a joke or not. We have no criterion at our disposal before our investigation has given us one. Linguistic usage is untrustworthy and itself needs to have its justification examined. In coming to our decision we can base ourselves on nothing but a certain 'feeling', which we may interpret as meaning that the decision is made in our judgement in accordance with particular criteria that are not yet accessible to our knowledge.²

This speaks for itself!

The schema given above illustrates how one has, in the context of discovery, a series of steps involving definition and redefinition of the terms '*Gx*' and '*Hx*' until *something like* the originally desired (3.2) is discovered. Nor is this, I believe, an inaccurate schematic representation of the general course of the activities of scientists as they attempt to achieve their ends. At the beginning there are certain tentative, partial definitions of terms, and certain tentative and more or less intuitive beliefs concerning correlations among these concepts. The scientist proceeds to check out these beliefs with which he begins. As he does this he comes to revise these beliefs. These revised beliefs lead him to revise in turn his definitions. Further aspects of his beliefs are now scrutinised, leading once more to revisions of both beliefs and definitions. At various places, the definitions may be revised in the interests of theoretical significance. And so on. Such is the progress of science toward the final product: a set of terms which are

¹ Trans. J. Strachey (New York, 1960), p. 15.

² Ibid. p. 61.

explicitly defined together with a set of laws arranged in a theoretical system, these laws being those that render the explicitly defined terms significant.¹

Such progress is no random groping. The scientist believes he can define the significant concept ' Gx '. These beliefs are not unreasonable. And it is these beliefs which serve to guide him:² What, then, are the origins of these beliefs which guide his activities and render them non-random?

First, the scientist takes hints from what he already has explicitly discovered in the area (e.g. (3.3) and (3.5)). Secondly, he can also be led to these beliefs by hints from the way the theoretical developments of the area are proceeding. Thirdly, there is pre-scientific usage, imprecise as it can be and with its vague borderline cases. Fourthly, there are our pre-scientific intuitions concerning correlations among the various factors, prejudiced as these intuitions often are.

In newly developing sciences, the last two factors are the more crucial. In well established sciences, it is the first two.

Guided by these hints and bits of knowledge, and above all by the desire to define a significant concept ' Gx ', the scientist indulges in various non-random activities during that process which has as its product the definition of the significant concept ' Gx '. *This non-random process might well be described as one in which the scientist gathers knowledge and adds indicators to the concept ' Gx ' until he has sufficient knowledge to produce an explicit definition of it.* Indicators are added; they close the area of indeterminacy of meaning; till at last there is no indeterminacy; one knows enough to be able to state an explicit definition of the desired significant concept; the product of the process is an explicit definition of the term to whose meaning we have been adding as we come to discover its exact significance. Or so goes the description.

This description is not inaccurate. Yet one must watch it carefully. For it contains two blurs.

In the first place, it blurs together two senses of the word 'meaning'. First, there is the meaning of a concept in the sense of 'that which is determined by the linguistic rules (of syntax, semantics, and definition)', i.e. its linguistically determined criteria of application. Second, there is the meaning of a concept in the sense of its significance, what might be referred to as its factually determined criteria of application. To speak of the 'meaning' of the term ' Gx ' changing may be to refer to either of these two senses of meaning.

¹ The ideal aimed at is 'process knowledge'. See n. 1, p. 43.

² From this it does not follow that there is any special 'logic of discovery'. I take this to be obvious.

The blur arises because, while one wishes to insist upon the distinction between these two meanings of 'meaning', one also wants to insist that questions of meaning (= significance) determine what meaning (= linguistic meaning) is given to ' Gx '.

In the second place, the description blurs what the scientist actually knows with what he hopes to discover. All that he knows to begin with are certain concepts, namely, ' Hx ', ' F_1x ', etc., and certain laws about these. He does not yet know that there is a law of the form (3.2). Nor does he yet have a definition of the term ' Gx ' which occurs in the law of the form (3.2) which he hopes to discover. The scientist *hopes to discover the significant concept ' Gx '*. To speak about the process in terms of 'adding indicators' to a concept may be to refer either to the redefining of a term already in the scientist's vocabulary, or to the search for the definition of that concept which the scientist hopes will be the product of his activities.

This second blur blends into the first. The scientist seeks to define ' Gx '. But only because he believes that the concept will in fact have a certain significance: he believes that there is a law of the form (3.2) which can be discovered. So to say that he expects to discover the definition of ' Gx ' amounts to saying that he expects to discover laws of a certain form.

This second blur arises in a way similar to that in which the first rose out of the description. One wishes to distinguish on the one hand, redefining a concept one already has and, on the other hand, hoping to define a concept with a certain significance. But one also wants to insist that both the redefinition and the scientist's hopes are determined by (his beliefs about) what the laws are, by the meaningfulness (= significance) of the concepts involved.

One might also note that a third meaning of 'meaning' tends to enter the description also. In this sense of meaning, the meaning of something is defined in terms of how one reacts to it. Thus, the cross is meaningful to a Christian in a way in which it is not meaningful to a non-Christian. So also, ' Gx ' is meaningful, in this sense, to the scientist, because he hopes to discover a definition of it such that (3.2) is true. But another mark, in which he takes no (professional) interest is, in this sense, meaningless to him.

Thus, the noted description of the process of the growth of knowledge blurs stipulation and fact (law), knowledge and anticipation. *If one does not have the relevant distinctions firm in one's mind, it is clear that one could easily move from the above description to the reduction sentence analysis of section I.*

Begin with noticing how ' Gx ' appears at the beginning of the process. It is undefined and uninterpreted. (3.2) is a universally quantified sentence

which relates this undefined and uninterpreted predicate to ' Hx ', that is, by virtue of the definition (3.1), to the ' H_1x '. The known truth (3.5) relates it again to ' F_1x ' and ' F_2x '. (3.2) guarantees that the presence of H_1 , H_2 or H_3 is each a sufficient condition for applying ' Gx ' to the individual; (3.2) and (3.5) guarantee that F_1 and F_2 are necessary conditions. *The general sentences which relate ' Gx ' to these other concepts,¹ that is, the axiomatics of ' Gx ', ' Hx ' and ' Fx ' provide necessary conditions, and sufficient conditions for applying the undefined and uninterpreted predicate ' Gx ',² though not conditions which are both necessary and sufficient.³*

Now recall that a reduction sentence is simply a general sentence relating an undefined and uninterpreted predicate to concepts whose meaning has already been determined, and notice Hempel:

A set of reduction sentences for a concept Q lays down a necessary condition for Q and a sufficient one; but, in general, the two are not identical. A definition of Q , on the other hand, specifies, in the definiens, a condition which is both necessary and sufficient for Q .⁴

Further comment is not necessary.

Again, in the case of ' Gx ' there is a certain area of indeterminacy, an area in which ' Gx ' cannot be applied. This is the case with those objects which falsify (3.7) and therefore also (3.6). Since (3.7) is false, the necessary conditions for applying ' Gx ' might be fulfilled by an object (' F_1x ' and ' F_2x ' might be true of it) while no sufficient condition is fulfilled (' Hx ' would be false of the object). The area of indeterminacy can be narrowed by adding further necessary conditions appropriate to the anomalous cases in such a fashion that if the necessary conditions are fulfilled then so are the sufficient; thus, one comes to add ' F_3x ' as a necessary condition for ' Gx '.

Once more notice what Hempel says about reduction sentences:

The indeterminacy in the meaning of a term introduced by a reduction sentence may be decreased by laying down additional reduction sentences for it which refer to different test conditions.⁵

This suffices to illustrate how a blur between meaning (linguistic meaning) and significance (axiomatics) could contribute to a reduction sentence analysis of the process we are considering.

' Gx ' is to begin with an undefined and uninterpreted predicate; only at the end of the process does it lose this status. The reason it is undefined

¹ Not simply: predicates!

² Not: concept!

³ We might say that the general sentences of the axiomatics of ' Gx ' 'implicitly define' the predicate ' Gx '. But such terminology, however fashionable, is misleading; see 'Implicit Definition Once Again', loc. cit., and 'The Notion of Logical Necessity in The Later Philosophy of Rudolph Carnap', ch. III.

⁴ C. G. Hempel, *Fundamentals of Concept Formation in Empirical Science*, p. 27.

⁵ Ibid. p. 27.

and uninterpreted is because *the scientist is as yet ignorant of the law he hopes to discover*. What the scientist knows is (3.5). What he hopes to discover is a law of *the form* (3.2), where, at the beginning of the process, (3.2) is not a law at all, but merely (because ' Gx ' is uninterpreted) the form of one. Of course, there *would* be a law of the appropriate form if (3.6) *were* true. But (we are supposing) (3.6) is in fact false. So the scientist does not have what he hopes to have, namely, knowledge of a law of the form (3.2).¹

In short, to indicate that at the beginning of the process ' Gx ' is undefined and uninterpreted is to indicate no more than that the scientist is as yet ignorant of, and that he hopes to come to know, a law of a certain form. The point is, the definitionalist's thesis is a thesis about *what the scientist knows*. It is not a thesis about how one might represent *the scientist's psychological states* of ignorance and hope. To suggest that the logical analysis of what the scientist knows ought to include features which indicate the psychological state of the scientist, is to do no more than confuse logic and psychology. The undefined and uninterpreted nature of ' Gx ' simply indicates certain psychological states; it represents, not a *concept*, but the *hope of a concept*. It therefore has no place in the logical analysis of what the scientist knows. As for what the scientist *knows* (at the beginning: (3.5); at the end: (3.9) = (3.13)), the definitionalist's thesis stands.

On the other hand, were one to blur the distinction between what is known and the hopes to come to know, or, more generally, between logic and psychology,² then one would find it hard to resist the move to reduction

¹ Let me point out that this is not the only way to look at the process. Another way is this: At the beginning of the process the scientist accepts the following law statement hypothetically:

$$(a) \quad (\exists g)[(x) (gx \equiv Hx)]$$

(a) corresponds in position to (3.2). ' g ' is, of course, a predicate variable, not an undefined, uninterpreted descriptive constant, as ' G ' in (3.2). The process is then viewed as a search for some predicate ' G ' which will verify (a); that is, as a search for some predicate ' G ' such that the law statement

$$(b) \quad (x)[Gx \equiv Hx]$$

is true. If the search is successful, if there is such a predicate, then (a) will be verified. For, (a) will follow logically from (b). The discovery of the law (b) will end the search, for the original hypothesis (a) will have been verified.

Nor, of course, is it an accident that (b) looks like (3.2). Hoping to verify (a) differs from hoping to discover a law of the form (3.2) by only the trivial factor of an existential generalisation. This shows that the two ways of looking at the process amount to the same.

The justification for taking the process as one which begins with an undefined and uninterpreted predicate, is simple. So interpreting it serves to illuminate certain roots of the reduction sentence idea.

With respect to the connection between (3.2) and (a), it is interesting to compare the comments on proper names such as 'Cicero' in N. L. Wilson, 'Substances without Substrata', *Review of Metaphysics*, 12, (1958-9), 521-39.

² For other aspects of this same confusion, see 'The Notion of Logical Necessity in the Later Philosophy of Rudolph Carnap', *passim*, and esp. ch. V.

sentences, as involving undefined and uninterpreted predicates, and as being replaced by an explicit definition as knowledge grows.

Let us once more look at Hempel:

... there must be 'operational' criteria of application for ... terms, i.e., criteria expressible in terms of observables. Reduction sentences make it possible to formulate such criteria. But precisely in the case of theoretically fruitful concepts, *we want to permit*, and indeed *count on*, the possibility that they may enter into further general principles, which will connect them with additional variables and will thus provide new criteria of application for them. We would deprive ourselves of these potentialities if we insisted on introducing the technical concepts of science by full definition in terms of observables.¹

Present interest is in the two italicized phrases. (a) As for 'permitting' concepts to have new criteria, *once these are known*, the definitionalist can adequately account for these things in terms of, at least, significance, and, perhaps also, depending on the specific case, redefinition. This we have already seen.² (b) As for what we hope to do, that is, as for our 'counting upon' adding criteria, it is sufficient to indicate three things: (1) all scientific concepts can be expressed in the language (either by introducing primitive predicates referring to observable properties, or by defining as one pleases or as one needs on the basis of these), from which it follows that all possible scientific laws, and more specifically, any law the scientist might discover, can be expressed in the language; (2) redefinition of terms already in the language is always possible; (3) given the possibilities of (1) and (2) all that remains of our 'counting upon' is certain psychological attitudes whose proper location is in the context of discovery, and which are therefore irrelevant to the logic of the situation, and in particular to the definitionalist's thesis.

This suffices for Hempel.

I have said that since (3.6) is false, the scientist does not have what he hopes to have, namely, knowledge of a law of the form (3.2). On the other hand, it will not do to say that our scientist knows nothing at all. (3.6) we supposed to hold in all except certain unusual and anomalous cases. What this supposition amounts to, is the assumption that (3.6), though it does not contain the 'whole truth', nonetheless is 'partially true'; that, in other words, (3.6) is a piece of 'imperfect knowledge'.³ (3.6) omits some relevant variables. This is why it has 'exceptions', a typical mark of imperfect knowledge. The omission of relevant variables makes (3.6) literally false. Such imperfection the scientist attempts to overcome. One of the omitted variables is discovered to be F_3 . Adding this factor, as in (3.8), removes part of the imperfection; the law statement (3.8) of the form (3.2)

¹ Hempel, *Fundamentals of Concept Formation in Empirical Science*, p. 29; italics added.

² See Section II of the present paper.

³ See n. 1, p. 43.

has fewer 'exceptions', is not falsified by so many instances, as the earlier formulation (3.6). But it too (we supposed) turned out to be false, there remained omitted relevant variables for which the scientist continued to search. The scientist works in such a fashion as to remove all imperfection from his knowledge. He continues until he has discovered general statements which admit no 'exceptions', that is, which are true.¹

Of course, what he discovers may not turn out to be what he had hoped or expected to discover. In our schema, such a dilemma appeared at (3.9). The scientist had two alternatives: (a) leave the axiomatics of ' Gx ' alone, that is, leave (3.2) and the definition (3.1) of ' Hx ' alone, and continue to search for a concept which will as a matter of empirical fact satisfy this grammar; or (b) change the axiomatics of ' Gx ', that is, alter the definition of ' Hx ' to ' H^*x ', and settle for the concept defined by (3.12) and which as a matter of empirical fact satisfies the grammar of (3.13).

We have now seen how one might move from certain features of the process of discovery to the reduction sentence analysis. We have seen how such a move depends upon blurring certain crucial distinctions. Such a move and such blurs are to be found in a paper of H. Putnam.²

Putnam indicates the process sketched above, and describes it as I suggested it might not unreasonably be described. He then suggests a reduction sentence analysis of it.

Our assumption or hope is that more precise indicators will be found for 'anger' (or for the various kinds of anger) until eventually an actual definition is possible; but the term 'anger' . . . is meaningful *now*, it does not merely become meaningful when it becomes possible to define it . . .³

Putnam insists that, of course we shall eventually define 'anger'. However, that does not mean the definitionalist's thesis stands; this is the force of Putnam's 'but'. It will not stand because it cannot account for the fact that 'anger' is meaningful *now*, before the explicit definition has been stated. We just saw, though, that the definitionalist need not disagree. 'Anger', he maintains, has not acquired meaning in the sense that a complete set of necessary and sufficient conditions for its application have not yet been laid down by linguistic stipulations. On the other hand, the definitionalist also maintains that: (1) There are *some* linguistic

¹ This is a necessary but not, of course, a sufficient condition for the law to conform to the ideal of scientific explanation. See n.1, p. 43.

² H. Putnam, 'Psychological Concepts, Explication, and Ordinary Language', *The Journal of Philosophy*, 54, (1957), 94-100. This paper of Putnam's contains two arguments against the definitionalist's thesis besides the one discussed in the present paper. I have examined these other arguments elsewhere. See 'The Notion of Logical Necessity in the Later Philosophy of Rudolph Carnap', ch. IV, sec. (4.3).

³ Putnam, op. cit. p. 99; his italics.

conventions governing the use of 'anger', i.e. there are definitions for certain terms 'anger₁', 'anger₂', etc. However, English, and scientists (unlike logical analysts) do not mark these with subscripts. (2) There is a certain amount of factual knowledge concerning lawful correlations among anger₁, anger₂, etc. (3) There is also a certain amount of factual knowledge as to how these concepts correlate with other concepts. (4) There are expectations on the part of the scientist that he shall be able to define the term 'anger'. Involved in these expectations are the expectations (4a) that the definition of 'anger' will involve the already defined 'anger₁', 'anger₂', etc., and (4b) that anger, when the term is defined, will be discovered to be a significant concept, and, indeed, (4c) that this significance will be of a more or less specific sort, the sort expected being determined by the background knowledge and its theoretical structure. Points (1)-(4) are more than sufficient to justify any assertion that 'anger' is meaningful now.

However, Putnam maintains that the process still cannot be one of definition and redefinition (and of searching for definitions). For,

When I discover that one of the indicators I have been using is a poor indicator (has declining correlation with the new indicators), I do not in fact say, 'Well, this person has condition C because that's how I define it'; rather I say, 'This person does not have the condition C, for *I was mistaken* in regarding this as an indicator.'¹

The preceding discussion has shown us that the definitionalist will have no problem in analysing this scientist's behaviour. The scientist hopes to define a significant concept. He expects certain things to enter into its definition. This expectation amounts to the belief that certain correlations exist and the suspicion that others exist. The hope amounts to the hope of discovering still other laws. This hope is based on yet other hints and suspicions. Sometimes these do not all check out against the facts as the scientist continues to acquire new knowledge during the process of discovery. Little wonder, then, that he will say at times that an indicator *he had believed* would enter into the definition of the significant concept he hopes to define, is a factor *he no longer believes* will enter into the hoped for definition.

It appears to me that we can best do justice to the attitude expressed in this way of speaking ('I was mistaken') if we say that by 'anger' I *now* mean the micro-state; that I regard the indicators not as defining characteristics but as symptoms; and that I anticipate that this list of symptoms will be progressively added to

¹ Ibid. p. 99; his italics. One might note how naively Putnam interprets the definitionalist's position. He seems to suggest that redefinition can occur only when the scientist specifically announces he is redefining a word. But why would anyone ever want to hold such a view? Certainly, the definitionalist is not committed to it. It is always easy to strike at straw men!

and revised so as to eventually determine an underlying condition (the micro-state) which causally explains the symptoms.¹

Now, *I* certainly don't mean the micro-state when I speak about a person being angry. Possibly some physiology-minded psychologists have meant that. But there is at least one person who means the micro-state, and that is Putnam. Or so he tells us. Let us take him at his word. The question is, can the definitionalist account for Putnam's meaning what he claims to mean? For the definitionalist to make his case go, he must show what it is for Putnam to *mean* a micro-state which has not yet been isolated.

The definitionalist's analysis of how Putnam means the not-yet-isolated micro-state is this: Putnam hopes to define, explicitly, and in the future, a concept 'anger' which will refer to a certain kind of physiological state. This kind of physiological state will be such that a person has it if and only if he also displays certain of the behavioural patterns commonly associated with anger. Such a correlation would ordinarily lead to a definition of a concept in terms of *both* the behavioural patterns *and* the physiological states.² However, Putnam also has certain (not unreasonable) expectations about the future theoretical development of psychology and believes that the laws which involve the micro-states will be of greater theoretical significance than those involving behaviour. These expectations lead Putnam to seek a concept which refers to the micro-state alone, rather than to both the micro- and behavioural states.³

In short, against the definitionalist, Putnam has no case.

University of Toronto

¹ Putnam, p. 99; his italics.

² See section II of the present paper.

³ Putnam is anticipating a definition of 'anger'. Compare G. Bergmann on the anticipation of the definition of 'mass' in classical mechanics; see his *Philosophy of Science*, pp. 90-91. See also E. Mach, *The Science of Mechanics* (Lasalle, Illinois, 1960), pp. 237 ff., 264 ff., 312 ff.

Discussions

CONFIRMATION STILL WITHOUT PARADOXES

Recently Howard Kahane, in a discussion in this *Journal* [3], has alleged that the solution of the confirmation paradoxes which I advanced in my 'Confirmation Without Paradoxes' [1] is unsatisfactory. As the title of this reply suggests, what is to be argued here is that Kahane's objections, not the proposal to remove the paradoxes by means of an eliminative confirmation system, are unsuccessful. Hence the following arguments are elaborations or extensions of the position taken in my original discussion. The objections advanced by Kahane can be summarised as claims of inadequacy in the proposed solution of the confirmation paradoxes in five respects: (1) the interrelations of necessary and sufficient condition statements; (2) the determination of the characteristics to be considered as members of the set of possible conditioning properties of a conditioned property; (3) the confirmation of disjunctive hypotheses; (4) the relation of confirmation to the denial of an hypothesis; (5) the status of 'd' instances as confirming instances. These will be considered in the order listed.

The primary objections Kahane raises concern the consequences of various equivalences for eliminative confirmation. As was clearly indicated in 'Confirmation Without Paradoxes', the following sample hypotheses are all equivalent to one another, and logically so ([1], pp. 178, 180-1):

1. $(x) \cdot Fx \supset Gx$
2. $(x) \cdot \sim Gx \supset \sim Fx$
- 1N. G is a necessary condition of F
- 1S. F is a sufficient condition of G
- 2N. $\sim F$ is a necessary condition of $\sim G$
- 2S. $\sim G$ is a sufficient condition of $\sim F$

It was not there held, as Kahane seems to think, that 1N is equivalent to 'G is a sufficient condition of F' ([3], p. 54); it was implicitly indicated that these two are converses of one another. It was also quite clearly indicated ([1], pp. 180-1) that whatever confirms any one of these several equivalent hypotheses confirms all the others also. Thus the problems, if any, which arise in regard to these several equivalent ways of reading 1 or 2 must concern what can count as evidence for both of them and for any of these ways of reading them. Not only is this so, but it is also the case that Kahane has neglected a central element in eliminative confirmation in attempting to raise difficulties here. Thus to set matters straight it is necessary once more to introduce the basic elements of eliminative confirmation.

Following von Wright ([6], pp. 127-30), these basic elements were formulated in terms of sets of characteristics which are possible conditioning characteristics of some other one and the elimination of members of such sets. On von Wright's

formulation, such a set is either a set of possible necessary conditions or a set of possible sufficient conditions. Hence the set designated A' was that of possible necessary conditions of F , that designated A'' possible sufficient conditions of G . Certain requirements were made to insure that these initial evidence sets provided a minimal basis for eliminative confirmation while not permitting the confirmation of copiously or vacuously true 'hypotheses'. It was, thus, required that every member of A' be known to be present with F in at least one instance and absent with F in another; the same was required for members of A'' in relation to G . To insure that the members of these two sets were in fact *possible* necessary conditions of F for A' and sufficient conditions of G for A'' , it was lastly required that there be no known instance in which a member of A' was absent while F was present or in which a member of A'' was present while G was absent ([1], pp. 179-80). If 1, say, is a possible hypothesis on the basis of either initial evidence set here, then its eliminative confirmation is provided by eliminating other members of that set than the one included in 1, i.e. by eliminating all but G from A' or all but F from A'' . (Thus to say that 1 is a possible hypothesis here is to say that G is a member of A' , F a member of A'' .) Since by its specification A' contains possible necessary conditions of F , elimination based on it is possible only by instances which have F but lack some previously uneliminated member of A' . Similarly, since A'' by its specification contains possible sufficient conditions of G , elimination based on it is possible only by instances which lack G but have some previously uneliminated member of A'' .

It is, then, a direct and obvious consequence of the schemata of eliminative confirmation that eliminative confirmation of 1 or 2 on the basis of A' must be by elimination of possible necessary conditions of F . It is similarly a direct and obvious consequence of these schemata that eliminative confirmation of 1 or 2 on the basis of A'' must be by elimination of possible sufficient conditions of G . This limitation by way of the initial evidence of what can be an eliminative confirmation of an hypothesis has nothing essentially to do with the way(s) in which we read the hypothesis. It is 1 and 2 that are confirmed, whether they be read as 1N, 1S, 2N, or 2S. Kahane's error here is the complete omission of consideration of this role of the initial evidence set. The error is compounded by his failure to appreciate the relations which do in fact obtain between these evidence sets and the various ways of reading 1 and 2. Given that the initial evidence is A' , it is 'natural' to read 1 as 1N because and only because this initial evidence includes the assumption that F has a necessary condition. But this 'naturalness' hardly mandates such a reading. Saying that F *has* a necessary condition is all one with saying that it *is* a sufficient condition, and a sufficient condition of precisely that characteristic which is F 's necessary condition. Similarly, given that the initial evidence is A'' , it is 'natural' to read 1 as 1S since this initial evidence includes the assumption that G has a sufficient condition. But since to say that G *has* a sufficient condition is all one with saying that it *is* a necessary condition, it is equally feasible to read 1 as 1N even given the initial evidence A'' . *Once more, what confirms 1 confirms it no matter how it is read; the same goes for 2.*

The foregoing remarks lead immediately to the removal of another of Kahane's objections with regard to the interrelations of various ways of reading 1. This attacks my claim that, given that 1 is read as 1S but the initial evidence is A' ,

we are not as it were searching for the sufficient conditions of G ; rather we are looking for that property of *which* F is a sufficient condition ([1], p. 181).

Kahane supposes that this introduces an element of psychologism ([3], p. 55) and responds that the idea of searching for something has no place in confirmation theories. He is correct in holding that psychologism has no place in confirmation theories but mistaken in supposing that that is what is here introduced. It seems he simply neglects the 'as it were' qualification in the just-quoted passage; at a minimum he seriously misreads it. What is being claimed here can easily be reformulated to remove any possible ambiguity as follows: given that 1 is read as $1S$ and that A' is the initial evidence, the hypothesis to be confirmed by eliminative instances is an hypothesis stating that F is a sufficient condition of some second characteristic. Finally, to attempt to forestall further misunderstandings let me rephrase the claim ([1], p. 179) that whether A' or A'' is chosen as the initial evidence is roughly reciprocally dependent upon whether 1 is read as $1N$ or $1S$. This amounts only to the claim that there is a certain 'naturalness' in reading 1 as $1N$ given A' and conversely, as $1S$ given A'' and conversely.

The last difficulty which Kahane attempts to raise in relation to the several equivalent ways of reading 1 and 2 concerns the question of whether or not 1 can simultaneously be confirmed by an instance which is eliminative with regard to A' and by one which is eliminative with regard to A'' ([3], pp. 53-54). To use his example, suppose we have an instance such as $FG \sim Ma$ and one such as $\sim F \sim Gmd$; do not both confirm or does not either one confirm 1 equally well? The answer is that both can confirm 1 and might confirm it approximately equally well, but that certain conditions must be met in each case. First, both can confirm 1 only if 1 is being confirmed on the basis of initial evidence which includes the set A' and also the set A'' . This is so since, although the two instances involve the same three characteristics, they do not yield the same information. Instance a here provides the information that M is not a necessary condition of F . Instance d here provides the information that M is not a sufficient condition of G . There is no reason why both initial evidence sets should not be used, and in a more developed eliminative confirmation theory than that used in 'Confirmation Without Paradoxes' there are good reasons for introducing both. Briefly, this has to do with the confirmation of hypotheses as statements of simplest sufficient or necessary conditions. Space limitations here, as in the original presentation, rule out the inclusion of such elaboration of the confirmation patterns of eliminative confirmation. Secondly, both instances here might confirm 1 approximately equally well given that the number of remaining possible necessary conditions of F after the observance of a is the same as the number of remaining possible sufficient conditions of G after the observance of d , and that the eliminative effect of the two instances is the same, i.e. that a eliminates as many possible necessary conditions as d eliminates sufficient ones. It is necessary to say approximately equally well here since if the total initial evidence is A' and A'' , the observance of either of the two instances at issue will reduce the number of remaining possible conditioning properties but the observance of the subsequent instance will reduce this number further and proportionately somewhat more.

The second major difficulty which Kahane attempts to raise is that of the determination of the membership of the initial evidence sets. This is raised in

two forms. First, Kahane objects that eliminative confirmation can be said genuinely to support an hypothesis only if genuine competitors are eliminated ([3], p. 53). Secondly, he objects that the confirmation theory permits the confirmation of 'all Martians are pink' given that we have observed only one pink Martian ([3], p. 55). These two objections are related since on Kahane's own view it would not be correct to say that we had genuine confirmation of the pink Martian hypothesis unless the characteristics included in the initial evidence on the basis of observation of the one known inhabitant of Mars were all genuine competitors. The difficulty in the demand that all the members of the initial evidence set be genuine competitors lies in providing criteria for this which do not amount to the requirement that we know in advance that the hypothesis at issue is correct. More specifically, it cannot be required that all the genuine competitors be known to be members of the initial evidence set(s) unless the intention is to resurrect a classical 'quest of certainty' version of eliminative induction. It does not seem that Kahane is seeking such a system and no such system was involved in the resolution of the paradoxes here at issue. Secondly, the criteria must permit formulation of initial evidence sets without reference to previously confirmed hypotheses in areas where investigation is just getting under way. Thirdly, the criteria must nonetheless be formulated in such a way that advantage can be taken of our knowledge that other generalisations are well confirmed where this is available and relevant.

It is a consequence of the foregoing considerations that the primary criteria for determining membership in the sets of characteristics constituting the initial evidence are those derived directly from the concepts of necessary and sufficient conditions. It was these criteria, and these alone, which were explicitly introduced above in the specification of A' and A'' . In bare contexts, to borrow and modify Mackie's terminology ([4], p. 266), these criteria have the consequence that all those characteristics which occur in the initial instance in which the characteristics of the hypothesis at issue occur and also do not occur in that instance in which the characteristics of the hypothesis at issue are absent are equally qualified as members of the initial evidence. But not all contexts are devoid of background information, and this can well serve to indicate which of the multitude of characteristics qualified by the minimum criteria for membership in the initial evidence set are plausibly possible alternatives. It is, it appears, this which is the root of the cogency of Kahane's example of the pink Martian. On the minimum basis, the observation of one pink Martian, coupled with some observation of a non-pink non-Martian, is sufficient for the determination of an initial evidence set. Kahane, however, here neglects to note that the number of characteristics so qualified will be quite high and the kinds of characteristics so qualified quite unrestricted. What makes the example a forceful one, despite this, is the unnoted background knowledge we all bring to this sort of case. We all know perfectly well that there are numerous species that do not exhibit uniformity of colour. Moreover, we know that in those species where such uniformity is found it is a consequence of something other than mere location. Hence there are at least two counts which serve to suggest quite strongly not only that the hypothesis that all Martians are pink is false but also that the genuine competitors for being a sufficient condition of being pink do not include being a Martian. Where such background knowledge is absent, it would be

entirely appropriate to attempt to delimit the initial evidence, make it a manageable set, by looking for more pink Martians and then using the positive analogy among the characteristics they have as the initial evidence.

But suppose, it will still be urged, that eliminative confirmation is attained for the hypothesis that all Martians are pink and we then discover one that is orange? What then? What then but that we have a well-confirmed generalisation which has turned out to be false? Again, unless it is insisted that every application of eliminative patterns must yield successful results, or to put it more accurately, that eliminative procedures must yield certainty, such results should not be surprising. What a confirmation theory proposes to do, having forsworn this illusory goal, is to provide a basis for evaluating putative evidence for an hypothesis, no more. Such evaluation can include the development of a new initial evidence set of characteristics and confirmation of a new hypothesis concerning the same characteristic considered in the previously advanced hypothesis which is now known to be false. But it is also possible, if there is some basis for supposing that the original evidence set did not include all the genuine competitors, to expand that set even although no falsifications of the hypothesis at issue are known. Hence Kahane is in error in urging ([3], p. 55) that if we have confirmed 1 on the basis of initial evidence earlier available and then develop additional genuine competitors we cannot provide further confirmation of 1 by eliminating members of the new set. It should be noted in this connection that the degree of confirmation of the hypothesis involved will at least temporarily be decreased by the introduction of further genuine competitors.

The third objection which Kahane raises is that eliminative confirmation is applicable only to hypotheses formulated as conditionals and not to ones formulated as disjunctions ([3], p. 53). This is, as one would expect from the equivalences holding between conditionals and disjunctions, simply wrong. If 1 is rewritten as the equivalent disjunctions:

$$3. (x). \sim Fx \vee Gx$$

and it is supposed that F is the conditioned property, then eliminative confirmation of 3 is attained by ruling out genuine competitors of G as the conditioning property. But this can only be done by instances in which F and G are present while some previously uneliminated member of the initial set of possible conditioning properties is absent. That is, it can only be done by precisely those eliminative patterns applicable to the confirmation of 1 given that A' is the initial evidence. On the other hand, if G is the conditioned property, eliminative confirmation is attainable only by ruling out genuine competitors of F as the conditioning property. This can, however, only be done by instances in which F and G are both absent while some previously uneliminated member of the initial set of possible conditioning properties is present. That is, it can only be done by precisely those eliminative patterns applicable to the confirmation of 1 given that A'' is the initial evidence. But surely none of this is surprising. To eliminate a characteristic as a necessary condition of F is to falsify the universal conditional having F as antecedent and that characteristic as consequent, and this is *eo ipso* the falsification of the equivalent universal disjunction. Similar considerations hold for the elimination of characteristics as sufficient conditions of G. Consequently, a *b* instance, i.e. one which is $\sim FG$, can eliminate no genuine competitor of 3, just as it can not provide eliminative confirmation of 1 as

was originally argued ([1], p. 180). Kahane will have to give up either his intuition that *b* instances confirm 3 or his intuition that genuine competitors must be ruled out in eliminative confirmation. They are inconsistent. Indeed, the whole problem of the paradoxes of confirmation results from such inconsistent intuitions, and no more than this need be said regarding the weight of any appeal to them.

The fourth objection which Kahane proposes is that eliminative confirmation is deficient since, although it may be applicable to the confirmation of an hypothesis, it cannot provide confirmation of the denial of that hypothesis ([3], p. 53). This is both true and irrelevant. If 1, for example, is false or more precisely known to be false, confirmation is simply not involved. The falsity of 1 is entailed by any falsifying instance of it, and confirmation is not entailment. Nor should it be so.

The fifth objection which Kahane attempts to offer is that eliminative confirmation will either permit an hypothesis to receive support from many *a* instances and also many *d* ones or from few *a* instances and thereby *d* ones ([3], p. 56). An *a* instance is one in which the characteristics mentioned in the 'positive' statement of the hypothesis are both present (F and G for 1), a *d* instance one in which both these characteristics are absent (cf. [1], p. 178). Kahane holds that this is in either case unacceptable because confirmation by *d* instances is counterintuitive. But to hold that this is so is not merely to make the already discredited appeal to intuition; it is also just to misunderstand the 'before' and 'after' stages of any experiment. The information concerning the 'before' stage is information about a *d* instance, and is essential. Consider the simple case in which zinc is placed in a flask; before an acid is added no gas is given off; after one is added, bubbles appear. Again, consider the common hypothesis that if a match is to light it must be dry. Surely an important part of the supporting evidence for this is our knowledge of cases of wet matches which have not lit when struck. Only on pain of ruling out *all* such evidence for hypotheses can one say that *d* instances are not to be admitted as confirming ones. (For additional discussion of this problem, see [5] and [2].)

In sum, then, confirmation by means of eliminative patterns is not embarrassed by any failure to take account of logical equivalences among various hypothesis formulations. It does involve the requirement that elimination be elimination of genuine competitors. It is quite unembarrassed by the fact that it is not possible to confirm the denial of a generalisation. It is at no disadvantage whatever, but instead recognises an important element in confirming evidence for hypotheses, by allowing 'negative' instances to count as confirmations. There is, then, no basis in Kahane's arguments for supposing that the removal of the confirmation paradoxes which eliminative patterns provide is not completely successful.

WILLIAM H. BAUMER

State University of New York at Buffalo

REFERENCES

1. Baumer, William H., 'Confirmation Without Paradoxes', this *Journal*, 15 (1964-65), 177-95.
2. Baumer, William H., 'Invalidly Invalidating a Paradox', *Philosophical Quarterly* (Scots), 15 (1965), 350-2.

3. Kahane, Howard, 'Baumer on the Confirmation Paradoxes', this *Journal*, 18 (1967-68), 52-56.
4. Mackie, J. L., 'The Paradox of Confirmation', this *Journal*, 13 (1962-63), 265-77.
5. Sharpe, R. A., 'Validity and the Paradox of Confirmation', *Philosophical Quarterly* (Scots), 14 (1964), 170-3.
6. von Wright, Georg H., *The Logical Problem of Induction*, 2nd revised edition, (Oxford, 1957).

MILLER'S PARADOX OF INFORMATION

Miller¹ and Popper² have claimed that the 'straight rule' (Reichenbach's rule of induction) is inconsistent, on the basis of a 'paradox' formulated by Miller. Mackie³ has suggested that Miller's argument is fallacious, but his attempt to analyse the paradox in terms of the quantifiers involved at each stage of the derivation is essentially irrelevant. In a reply,⁴ Miller has rightly rejected this criticism. However, he has unfortunately further confused the issue by ignoring the central thesis of Mackie's note.

Mackie pointed out, quite correctly, that this paradox has nothing to do with the concept of probability. In fact, it arises quite trivially whenever a function is not well-defined. The use of a functional notation such as $F(x) = y$ generally involves the implicit assumption that F is a one-one or many-one mapping from the set X of all elements x into or onto the set Y of all elements y . In other words, it is assumed that F is not a one-many mapping. This means, in particular, that equivalent elements in the domain of the function (the set X) are mapped onto equivalent elements in the image set or range of the function (the set Y), i.e.

if	$F(x) = y$
and	$F(x') = y'$
then	$y = y'$
if	$\vdash x = x'$

(This would not follow if F were a 'multi-valued function', unless the first two equalities above referred specifically to a particular 'branch' of the multi-valued function. A multi-valued function can always be regarded as a family of related single-valued functions.)

For example, consider the 'numerator function', which maps the set of all rationals of the form a/b (where a and b are integers) onto the set of all integers, as follows:

$$N\left(\frac{a}{b}\right) = a \quad (1)$$

¹ This *Journal*, p. 59, May 1966.

² This *Journal*, p. 61, May 1966; p. 141, August 1966.

³ This *Journal*, p. 144, August 1966.

⁴ This *Journal*, p. 147, August 1966.

By the above definition of the function N , therefore,

$$N\left(\frac{2a}{2b}\right) = 2a \quad (2)$$

Now, since arithmetically

$$\vdash \quad \frac{a}{b} = \frac{2a}{2b} \quad (3)$$

it should follow that

$$N\left(\frac{a}{b}\right) = N\left(\frac{2a}{2b}\right) \quad (4)$$

if N is a well-defined function. However, by (1), (2), and (4), we would then have:

$$a = 2a, \text{ or } 1 = 2 \quad (5)$$

Clearly the function N is not well-defined by equation (1), since it maps arithmetically equivalent elements in the domain of the function (the set of all rationals) onto arithmetically non-equivalent elements in the image set or range of the function (the set of all integers). The ambiguity can be removed by simply defining:

$$N\left(\frac{ka}{kb}\right) = a, \text{ for all integers } k \quad (6)$$

where a and b have no common factor. Alternatively, $N(r/s)$ can be regarded as a multi-valued function with different branches:

$$N\left(\frac{r}{s}\right) = ka \text{ where } k = 1, 2, 3 \dots \quad (7)$$

Here $r/s = na/nb$; a and b have no common factor. The different branches can be distinguished by a subscript, i.e.

$$N_k\left(\frac{r}{s}\right) = ka \quad (8)$$

More generally, consider the set of all arithmetic propositional functions of the form:

$$x = ax + b \quad (9)$$

where a and b are any real numbers, including zero. Define a mapping from the set of propositional functions $\{x = ax + b\}$ onto the set of functions $\{ax + b\}$ as follows:

$$F(x = ax + b) = ax + b \quad (10)$$

From this definition of the mapping F , it follows that:

$$F\left(x = \frac{b}{1-a}\right) = \frac{b}{1-a} \quad (11)$$

Now, arithmetically,

$$\vdash \quad (x = ax + b) \equiv \left(x = \frac{b}{1-a}\right) \quad (12)$$

Hence, if F is a well-defined function,

$$F(x = ax + b) = F\left(x = \frac{b}{1-a}\right) \quad (13)$$

Hence, by (10), (11), and (13), we have:

$$ax + b = \frac{b}{1-a} \quad (14)$$

If it is not immediately clear that this result is paradoxical, take $a = -1$, $b = 1$. Condition (12) is satisfied, and the argument runs as follows:

$$F(x = 1 - x) = 1 - x \quad (10')$$

$$F(x = \frac{1}{2}) = \frac{1}{2} \quad (11')$$

$$F(x = 1 - x) = F(x = \frac{1}{2}) \quad (13')$$

Hence:

$$1 - x = \frac{1}{2}, \text{ or } x = \frac{1}{2} \quad (14')$$

The reason for this apparently paradoxical result is again that the function F is not well-defined. The domain of F is the set $\{x = ax + b\}$, and the range of F is the set $\{ax + b\}$. Since in the domain of F , the arithmetic propositional function $x = ax + b$ is logically equivalent to the arithmetic propositional function $x = (b/1-a)$, we must stipulate whether this propositional function is mapped by F onto $ax + b$ or onto $b/1-a$. Evidently $ax + b$ and $b/1-a$ are not in general equivalent elements of the image set of F : in the previous example $1-x \neq \frac{1}{2}$ unless $x = \frac{1}{2}$.

In Miller's paradox,¹ the logical probability P is defined as a mapping from a set of information statements $\{p(a) = r\}$ onto the set $\{r\}$, where ' $p(a) = r$ ' (abbreviated as ' E_r^a ') is interpreted as the statement that 'the (statistical) probability of event a , on the given chance set-up, equals r '.

Thus:

$$P(A, E_r^a) = r \quad (15a)$$

where A is the statement that event a occurs.

The statement A plays no role in Miller's deduction. Equation (15a) states formally that E_r^a is mapped onto r . The interpretation of this mapping is that the value r refers to the logical probability of a proposition corresponding to the event a . Therefore, suppressing the proposition A , equation (15a) may be written in the form:

$$P(p(a) = r) = r \quad (15b)$$

The reason for expressing equation (15a), which corresponds to Miller's equation (1), in the form (15b) is merely to emphasize the functional form of (15a), as opposed to the probabilistic interpretation imposed by Miller. Equation (15b) is not a reformulation or translation of Miller's notation and, indeed, Miller's notation can be substituted throughout the following argument with no alteration in the logic. It is then immediately evident that, *relative to the derivation of the*

¹ The following discussion of the paradox refers to Miller's note in this *Journal*, p. 59, May 1966. The numbered equations referred to are from this note.

paradox, Miller's notation is both redundant (with respect to the proposition A) and obscure (with respect to the substitution of ' E_r^a ' for the expression ' $p(a) = r$ ').

The paradox arises as in equations (10'), (11'), (13'), and (14') because P is not a well-defined mapping. Substituting $p(a)$ for x and writing $p(-a)$ explicitly as $1-p(a)$, we have:

$$P(p(a) = 1-p(a)) = 1-p(a) \quad (16)$$

(cf. Miller's equation (6))

$$P(p(a) = \frac{1}{2}) = \frac{1}{2} \quad (17)$$

(cf. Miller's equation (5))

$$P(p(a) = 1-p(a)) = P(p(a) = \frac{1}{2}) \quad (18)$$

(cf. Miller's equation (4))

Hence:

$$1-p(a) = \frac{1}{2}, \text{ or } p(a) = \frac{1}{2} \quad (19)$$

(cf. Miller's equations (7) and (8))

Miller's account is persuasive because his notation leads us quickly and deceptively through the steps of his proof. For his equation (1)

$$P(A, E_r^a) = r$$

which expresses the straight rule, there is a tendency to read the subscript of ' E_r^a ' to obtain the value of the function P . This habit seems innocuous. However, in conjunction with the precise criterion which the probability calculus gives us for identity of arguments (identity of E -expressions means logical equivalence), the function P becomes multi-valued and allows Miller to derive a contradiction. We can avoid this trap by recalling the original purpose of introducing the E -expressions: They were meant to inform us of the statistical probability of a . So, we can correct matters by interpreting Miller's (1) as stating that ' E_r^a ' designates an equation $p(a) = r$ which, if yielding a value of $p(a)$ by solution, assigns that value to P . In this case it is natural to define¹:

$$P\left(A, E_{r p(a)+s}^a\right) = \frac{s}{1-r} \quad (20)$$

¹ More generally, the straight rule can be regarded as a transformation rule from statements about $p(a)$ —the statistical probability of event a —to statements about $P(A)$ —the logical probability of the statement A that event a occurs. If $p(a)$ is known to satisfy the functional relationship $p(a) = f(p(a))$, then the straight rule should simply transfer this functional relationship to $P(A)$, i.e. $P(A) = f(P(A))$. In Miller's notation, this reads:

$$P(A, E_{f(p(a))}^a) = f(P(A, E_{f(p(a))}^a)), \text{ and not: } P(A, E_{f(p(a))}^a) = f(p(a)).$$

In the case of a *linear* function f (explicitly, $p(a) = rp(a) + s$, or $P(A) = rP(A) + s$), the equation $P(A, E_{r p(a)+s}^a) = rP(A, E_{r p(a)+s}^a) + s$ has the unique solution $P(A, E_{r p(a)+s}^a) = s/(1-r)$ (see equation (20)). Miller derives the paradox from the equation $P(A, E_{r p(a)+s}^a) = rp(a) + s$ (with $r = -1$, $s = 1$), which is consistent with the equation $P(A, E_{r p(a)+s}^a) = s/(1-r)$ (or $P(A, E_{r p(a)+s}^a) = rP(A, E_{r p(a)+s}^a) + s$) expressing the straight rule, if and only if $p(a) = s/(1-r)$.

Miller's equation (6):

$$P\left(A, E_{p(-a)}^a\right) = p(-a)$$

is then false and should read

$$P\left(A, E_{p(-a)}^a\right) = \frac{1}{2}$$

Evidently these two equations are only equivalent if $p(-a) = p(a) = \frac{1}{2}$.

Thus, the paradox disappears once the ambiguity in P is removed and the straight rule is protected from this line of attack.

JEFFREY BUB and MICHAEL RADNER
University of Minnesota

LENGTH CONTRACTION AND CLOCK SYNCHRONISATION: THE EMPIRICAL EQUIVALENCE OF THE EINSTEINIAN AND LORENTZIAN THEORIES

Professor Popper¹ says that Lorentz and Fitzgerald 'do *not* say that a rod *at rest in the ether* would *appear contracted* if measured with the rods of the moving system. (Rather, one might expect, by implication, that it would appear *expanded*.)' Popper then goes on to emphasise the mutual character of the Einstein contraction, and stresses that it is this mutual character of the contraction which leads to the abandonment of the ether.

While everything that Popper actually says seems to me perfectly correct, I should like to try to clarify the remark he makes in brackets in the passage I have quoted above, in order to clear up a possible misconception.

I shall restrict my considerations to a Lorentzian theory in which the velocity of light is held to be constant relative to something like the fixed stars, and where both length-contraction and clock-retardation are supposed to occur when bodies move relative to this framework. I want to explain and defend the following two statements:

(I) On such a theory the rod at rest in the ether *would appear expanded* if measured with the measuring rod of the moving system *provided* the clocks of the moving system are correctly synchronised according to a *Lorentzian prescription*.

(II) On the other hand such a theory predicts that the rod at rest in the ether *would appear contracted* if measured with the measuring rod of the moving system *provided* the clocks in the moving system are synchronised according to an *Einsteinian prescription*.

Here the Einsteinian prescription is the usual one, whereas the Lorentzian clock synchronisation prescription is implicit in the Lorentzian light postulate. For while Lorentz as far as I know never said anything very explicit on this point, it is clear that according to his light postulate, namely the constancy of the velocity of light relative to the ether, he cannot accept the Einsteinian method of clock synchronisation as giving correct results in systems moving relative to

¹ K. R. Popper, this *Journal*, 16 (1966), 332.

the ether. Only for a system of clocks stationary relative to the ether will a Lorentzian accept Einstein's method of clock synchronisation as correct. Hence the theoretically simplest Lorentzian clock synchronisation prescription would involve setting up a system of clocks in the ether, synchronising them according to Einstein's method, and then prescribing that two events occur simultaneously if and only if they occur simultaneously according to the ether clocks. This is rather impracticable. In practice the Lorentzian method would involve knowing the velocity of a given system relative to the ether and synchronising its clocks using light signals, but making due allowance for the postulated difference in the backwards and forwards relative velocity of light. For example: If the system has the velocity $\frac{1}{2}c$ relative to the ether, and if a light signal sent from one of the system's clocks at time t and travelling in the line of motion returns to it at time t plus four seconds, after having been reflected by another of the system's clocks, then the Lorentzian prescription is that we set the second clock so that the time at which it reflected the light was t plus three seconds or t plus one second depending on whether it was in front of or behind the first clock. This differs from the Einstein prescription of t plus two seconds, because for the Lorentzian the velocity of light relative to the system, instead of being treated as c in each direction, is postulated to be $c - \frac{1}{2}c = \frac{1}{2}c$ in one direction and $c + \frac{1}{2}c = 1\frac{1}{2}c$ in the other direction.

Notice that the Lorentzian methods of clock synchronisation presuppose that the ether reference frame can be empirically identified. That is why it is necessary to identify Lorentz's ether reference system with something like the fixed stars. That such an identification has to be by fiat is certainly an embarrassment to a Lorentzian (but this does not mean it need have no physical motivation and one has, in fairness, to bear in mind that it is a psychological fact about Lorentzians that they have not been able to understand how the Einstein-Minkowski alternative purports to explain the facts in question).

The proof of statement (I) above, which contains what is correct in Popper's expectation, is genuinely trivial. For on the Lorentzian method of clock synchronisation simultaneity is absolute: an ether observer and a moving observer compare the same measurements, and what is a contraction for one is necessarily an expansion for the other. So both Einstein and Lorentz would predict that a Lorentzian observer should observe an apparent expansion of the rod at rest in the ether.

Statement (II) gives us further assurance that we are not dealing here with an empirical difference between Einstein and Lorentz. Its proof is a moderately easy exercise in special relativity. (Indeed it is set as such as part of exercise 12, chapter 1, of Rindler's book.¹)

It is very important to appreciate just why clock synchronisation enters into the length measurement problem. To measure the length of a rod which is in motion relative to one's measuring rod, one has to note the position of the front end of the moving rod and the position of the rear end *at the same moment*. Thus if one's front clock is, on the basis of a certain criterion of synchrony, lagging a few seconds behind one's rear clock, then one will mistakenly mark off the far end of the rod one is passing at a later moment than one marks off its near end, and one will hence effectively reduce the rod's length by the amount the rod has

¹ W. Rindler, *Special Relativity*, 2nd edn. (Oliver and Boyd, 1966) p. 24. Cf. pp. 27-28.

travelled in the few seconds in question. But this is precisely the situation we are in from the Lorentzian point of view, if the clocks are synchronised by Einstein's method. For the relative velocity of light being less according to Lorentz in the forward direction than in the backward direction, the front clock, if synchronised by Einstein's method, will from Lorentz's point of view, lag behind the rear clock. If one actually computes the apparent foreshortening of the rod in the ether produced by this 'incorrect' clock synchronisation effect, one finds it is by a factor of $1 - (v^2/c^2)$. This is precisely that needed to convert the apparent expansion of $[1 - (v^2/c^2)]^{-1/2}$ which a Lorentzian observer would observe into the apparent contraction of $[1 - (v^2/c^2)]^{1/2}$ which an Einsteinian observer would observe. Both Einstein and Lorentz will agree on what observers who synchronise their clocks in specified ways will actually observe.

The same analysis applies with minor modifications to the measured rates of moving clocks. Here again it is essential to consider the way in which clock synchronisation enters into the measurement of the rate of a clock which is moving relative to the measuring system. Neglect of this is presumably responsible for Keswani's mistake in the argument which leads him to assert¹ 'It is sometimes said that there is no observable difference between the relativistic and the ether-theoretic point of view. This is incorrect.' He adds in a footnote 'I was not myself altogether clear on this point in part II of my paper'. If Keswani is understanding what I understand by the ether theory, then his earlier views, for which he admits there is strong authority in the literature, were surely the correct ones.

There is of course a profound philosophical difference between the Lorentzian and Einsteinian conceptions as Popper has emphasised, but the only area where their mathematical equivalence allows one any hope of this philosophical difference having an empirical cash-value seems to me to be in the cosmological generalisations of the respective theories. One would expect a cosmological generalisation of the Lorentz ether theory to take the postulated relationship between the ether and the fixed stars sufficiently seriously to yield a cosmological theory which was not equivalent to one starting from an Einsteinian interpretation of special relativity.

I have certainly never seen a clear statement of the theory I have been here attributing to Lorentz in his own writings. The best evidence which I have come across that Lorentz did eventually hold a theory empirically precisely equivalent to Einstein's is contained in the following comment² of P. Ehrenfest in the course of a lecture to an audience whose most distinguished member was Lorentz, the year after Ehrenfest had succeeded to Lorentz's chair at Leyden on Lorentz's recommendation :

Wir sehen also, dass hier die aetherlose Theorie von Einstein genau dasselbe verlangt, wie die Aethertheorie von Lorentz. Auf diesem Umstand beruht dann auch, dass nach der Einstein'schen Theorie ein Beobachter an den vor ihm laufenden Masstäben, Uhren etc. exact dieselben Contractionen, Gangänderungen u.s.w. beobachten muss, wie nach der Lorentz'schen Theorie. Und hier sei gleich allgemein bemerkt : Ganz principiell gibt es kein experimentum crucis zwischen diesen beiden Theorien.

JON DORLING

University of Sussex

¹ G. H. Keswani, this *Journal*, 17 (1966), 151.

² P. Ehrenfest, *Collected Scientific Papers* (North Holland, 1959) p. 320.

A NOTE ON CONSILIENCE

It is generally considered a consequence of what Whewell called 'consilience' that if two hypotheses, h_1 and h_2 , may both be treated as cases or exemplifications of a more general theory, T , then any evidence that raises the degree of confirmation for h_1 and therewith of that of T , must also raise the degree of confirmation for h_2 . Moreover, in these circumstances, if h_1 raises the degree of confirmation for T , it also raises that for h_2 . But there is a problem about how to elucidate this feature of consilience, especially since h_1 and h_2 will normally be logically independent of each other. For example, it may perhaps be tempting, at first sight, to suppose that the situation arises because if anything raises the degree of support for T it must do so for h_2 in accordance with some such consequence condition as Carnap's T 59-2d, viz. if $\vdash h_1 \rightarrow h_2$, then $c(h_1, e) \leq c(h_2, e)$.¹ But all that follows from this consequence condition here is that if e raises the degree of confirmation for T through raising it for h_1 then e raises the lower limit which h_2 's degree of confirmation must equal or surpass. It does not follow that e raises h_2 's actual degree of confirmation. h_2 may in fact have passed this limit in any case.

The object of the present note is just to point out that the feature of consilience in question is deducible from what I have elsewhere² called the Instantial Comparability Condition, viz. the principle that inequalities in degree of evidential support between two theories lead normally to inequalities in degree of evidential support for the applications of these theories and inequalities in the latter normally stem from inequalities in the former. We may write this principle as

For all E, E', S, S', U and U' , if

- (i) U is a universal statement, S is just a substitution-instance of U , and E is a conjunction of statements each of which predicates something of some specified element in the domain of U , other than the element referred to by S , and
- (ii) U' is a universal statement, S' is just a substitution-instance of U' , and E' is a conjunction of statements each of which predicates something of some specified element in the domain of U' other than the element referred to by S' ,

then $c(U, E) > c(U', E')$ if and only if $c(S, E) > c(S', E')$.

I shall not repeat here the informal arguments for the Instantial Comparability Condition and its important consequences for probabilistic theories of confirmation. But it is clear that the feature of consilience in question is deducible from the Instantial Comparability Condition, since the latter implies both

¹ *Logical Foundations of Probability*, 2nd edn., 1962, p. 317.

² In 'A Logic for Evidential Support', this *Journal* 17 (1966), 23 ff.

and

$$c(h_1, e \& e') > c(h_1, e) \rightarrow c(T, e \& e') > c(T, e)$$

and also

$$c(T, e \& e') > c(T, e) \rightarrow c(h_2, e \& e') > c(h_2, e)$$

$$c(T, h_1 \& e) > c(T, e) \rightarrow c(h_2, h_1 \& e) > (h_2, e)$$

where h_1 and h_2 are substitution-instances of T and e and e' are statements of the appropriate kind. Similarly, when we map the logical syntax of this type of confirmation-theory on to a generalisation¹ of the modal system S_4 , we obtain the corresponding theorems.

L. JONATHAN COHEN

The Queen's College, Oxford

¹ Certain corrections are necessary in the published version (ibid. pp. 114 ff.) of this system. Introduce propositional variables (represented by p , q or r) into the primitive symbolism. p and $\exists p(A)$ are to be wf. Axiom-schemata 29, 30, 31 and 32 are to be extended over propositional variables in the usual way. $c[H, E] < n$ is then to be redefined as an abbreviation for $\forall \Box^x ((E \rightarrow^c \Box^x H) \rightarrow \forall p (\Box^n p \rightarrow^d \Box^x p))$. Also delete 'or v is $\Box^x$ ' in axiom-schema 29; replace ' $\vdash B$ ' by ' $\vdash \forall \Box^x (B)$ ' in 33; and take E, F, G, H, I and J throughout to represent only wff in which no modal variable occurs freely.

Reviews

J. Hintikka and P. Suppes, eds: *Aspects of Inductive Logic* (Amsterdam, North Holland, 1966). Pp. 320.

In the second of his three interesting contributions to this volume ('Probabilistic Inference and the Concept of Total Evidence'), Patrick Suppes questions the legitimacy of that type of inference often called the 'statistical syllogism' and with it the importance of the concept of total evidence. Although Suppes has much of importance to say, his discussion is marred by his misrepresentation of the form of statistical syllogisms as commonly understood. Before considering the genuinely significant parts of his argument, therefore, it may be worthwhile spending some time in setting the record straight.

According to Suppes, statistical syllogisms have the following form:

$$(1) \frac{P(\text{hypothesis/evidence}) = r}{\text{evidence}} \\ P(\text{hypothesis}) = r$$

Suppes understands the P-function appearing in the major premiss and conclusion of (1) to be a measure of subjective probability expressing judgments as to fair betting quotients. He correctly observes that this type of argument violates the requirements of the calculus of probabilities save in the special situation where $P(\text{evidence/hypothesis}) = P(\text{evidence})$. Although Suppes subsequently acknowledges an 'inferential kernel inside the dubious statistical syllogism' (p. 59), he obviously thinks that the illegitimacy of (1) suffices as a warrant for refusing to take statistical syllogisms *per se* seriously.

After rereading Carl Hempel's extensive discussions of the statistical syllogism, I have not been able to find a single place where such arguments are cast in form (1). In Hempel's account, the major premiss of a statistical syllogism is a statistical hypothesis of the form $p(A, B) = r$ which asserts that the statistical probability of a B's being an A is r . This assertion is not to be confused with the claim that the conditional subjective probability of the hypothesis that a is an A given that a is a B is r . (This may be symbolised as $P(a \text{ is an A} / a \text{ is a B})$).

The second premiss of Hempel's syllogism is an assertion that a is a B. Finally, in some versions, the conclusion asserts that the subjective or inductive probability $P(a \text{ is an A})$ which determines fair betting rates on the truth of the hypothesis that a is an A is r . Hence, the form of the statistical syllogism may be restated as follows:

$$(2) \frac{p(A, B) = r}{a \text{ is a B}} \\ P(a \text{ is an A}) = r$$

Unlike in (1), the probability measures appearing in the major premiss and the conclusion of (2) are different and are defined over different sets of entities. Thus, whatever the difficulties which face the use of statistical syllogisms, they

cannot be dismissed as invalid on the grounds that the requirements of the calculus of probabilities are violated.

Suppes's misrepresentation of the form of the statistical syllogism may stem from his apparent adherence to the extreme subjectivist view advocated by Bruno De Finetti. According to De Finetti, the notion of statistical probability has obnoxious metaphysical implications which can be eliminated by strict adherence to a subjectivist interpretation of probability. Thus, ' $p(A,B) = r$ ' should be replaced by ' $Q(x \text{ is an } A/x \text{ is a } B) = r$ ' for arbitrary substitutions for the variable ' x '. Here the Q-function is a suitable subjective or inductive probability measure Q.

Although the expendability of the concept of statistical probability is a moot point, it may be conceded for the sake of the argument. This concession does not warrant Suppes's reconstruction of the statistical syllogism in subjectivist terms as having the form (1). To do so assumes that the Q-function representing the subjectivist reading of the statistical hypothesis is the same as the P-function which appears in the conclusion of a statistical syllogism. Suppes has adopted the most uncharitable subjectivist reconstruction of the statistical syllogism possible.

Suppes does finally turn to serious issues when he examines the 'inferential kernel inside the dubious statistical syllogism'. Without admitting that it is an appropriate subjectivist reading of the structure of the statistical syllogism, Suppes examines probabilistic arguments where the Q-function does differ from the P-function.

$$(3) \quad \frac{Q(h/e) = r}{P(h) = r}$$

Observe that arguments of type (3) are of critical interest even to those who insist on distinguishing the notion of statistical probability and reading statistical syllogisms as being of form (2). Hempel, for example, contends that the passage from the premisses of arguments of type (2) to the conclusion is licensed by an argument of form (3) where the hypothesis h asserts that a is an A and the 'evidence' e asserts that $P(A,B) = r$ and a is a B. Suppes's examination of arguments of this form is directly relevant, therefore, to discussions of the statistical syllogism even for those not committed to De Finetti's extreme reductionist posture.

It is well known that statistical syllogisms in particular and arguments of type (3) in general can lead to conflicting probability conclusions if they are not modified in certain ways. The usual recommendation is to require not only that the second premiss assert e but that e be the 'total' evidence in some sense or other. Thus form (3) ought to be modified as follows:

$$(4) \quad \frac{Q(h/e) = r}{P(h) = r}$$

e is the "total" evidence

Two questions immediately arise: (a) The P-function represents the probabilities to be used to determine fair betting quotients under the conditions actually obtaining. If the Q-function is a subjective probability function different from the P-function, it must represent probabilities which are used to determine fair

betting quotients under conditions different from those actually obtaining. What are these conditions? (b) How total must the 'total' evidence be?

Both of these questions are answered in a unitary way by Rudolf Carnap. The Q -function is understood to be a measure of confirmation in Carnap's sense. It represents probabilities which would be used to determine fair betting quotients were the agent in a situation where he had no empirical evidence. Hence, the premiss that $Q(h/e) = r$ is to be understood as meaning that were the agent to have e as his total evidence—i.e. were e all the empirical evidence available to the agent over his lifetime up to the moment under consideration—he should assign a probability of r to h . The second premiss in (4) is then read as affirming the antecedent of this subjunctive conditional and leading to the conclusion that $P(h) = r$.

Ignoring the real difficulties involved in constructing a suitable Carnapian confirmation function to use as the Q -function, Suppes points to another serious objection. The requirement of total evidence as understood by Carnap is inapplicable in any real life situation. Men lack the memories or time to dredge up and explicitly formulate all the evidence available to them (in some obscure sense of 'available') over their lifetime. Hence, either arguments of type (4) must be considered suspect or master Q -functions quite different from Carnapian confirmation functions must be found.

When considering rational determination of probability assignments to be made at a given time, Suppes adopts the first strategy. He thinks that conditionalisation relative to 'total' evidence *via* arguments of type (4) is utterly expendable in this context. He contends that the coherence requirements—which articulate the sensible maxim that a rational agent should avoid a Dutch book being made against him—take care of the contribution of the evidence to the determination of probability judgments made by an agent at a given time. Any statement admitted as evidence will, according to the coherence requirements, be assigned a probability of r . Thus, the evidence available at time t is distinguished by the extreme probability value assigned to evidence statements at t . Hence, '... it is automatic that if a person is asked for the probability of an event at a given time it will follow from the conditions of coherence on all of his beliefs at that time that the probability he assigns to the event automatically takes into account the total evidence that he believes has relevance to the occurrence of the event' (p. 59).

While dismissing conditionalisation in connection with rational probability judgment at a given time, Suppes takes it seriously in connection with changes in probability judgment over time. If the master function Q expresses probability judgments made at time t and relative to the evidence at t and e is the total *new* evidence acquired between t and t' , then the probability function P used at t' should, according to conditionalisation, be related to Q according to the equation $P(h) = Q(h/e)$.

Observe, however, that although conditionalisation so understood avoids the troubles confronting the Carnapian interpretation of the master Q -function, it does so by becoming a principle of normative learning theory. The relation between such principles and learning theory as a branch of empirical psychology is roughly the same as the relation between a theory of the behaviour of Volkswagens and a system of recommendations as to how they should be modified

in order to behave in a certain ideal fashion. It is clear that recommendations regarding the ideal Volkswagen are worthwhile only in so far as they can be implemented at least to a reasonable degree of approximation; and this presupposes a good deal of factual information about automotive mechanics. Similarly, any design for an ideal learner, if it is to be applicable, presupposes a good deal of information about human learning. The central lesson which Suppes attempts to inculcate in this essay as well as in his first contribution to this volume is that students of rational inductive inference and behaviour ought to focus their attention on learning theory.

To illustrate his thesis, Suppes cites some of the inadequacies which conditionalisation displays as a principle of learning from experience. In many types of situations, human beings lack both the memories and computational abilities required to conform to conditionalisation. Conditionalisation is impotent to handle conceptual innovation. It does not even begin to face the task of providing an account of how information is or ought to be selected from the signals pressing in on the human organism from his environment. Finally, it does not indicate how probability judgments change or ought to change in response to sensory stimulation which does not result in the acquisition of new evidence. This list of inadequacies can only be remedied, Suppes maintains, by looking to psychology.

Suppes's argument does carry considerable weight provided that theories of rational inference are intended to be systems of norms regarding the appropriate design for an ideal learner. However, this is not the only normative application which theories of rational inference have been intended to have or even, for many writers, a primary one. Philosophical interest in such theories often stems from an interest in determining conditions under which beliefs and decisions can be *justified*.

Suppose that at time t' Smith and Jones are contemplating some joint venture and are concerned to determine the probabilities to be distributed among some set of propositions for gauging risks. Smith uses a coherent probability function to assign $P(h)$ to h and Jones uses a coherent probability function $P^*(h)$. On such an occasion, Smith will be obliged to justify use of his coherent probability function to Jones if he is legitimately to recommend that Jones adjust his opinions to conform to Smith's.

Smith might try to offer some details of his biography. At time t shortly before t' he might have been using a probability function Q . Between t and t' , he might have acquired evidence e . If $Q(h/e) = P(h)$, he might attempt to justify his use of the P -function at t' to Jones by means of a conditionalising argument.

Will this argument constitute a successful justification? Not if Smith has merely succeeded in supplying some biographical details. Jones might admit e into the evidence as does Smith (although he may have acquired e at some different time than Smith). He might even agree with Smith about the 'background information'—i.e. about that portion of the 'total' evidence at t' which is not explicitly cited in e . Yet, he might maintain that were the total evidence at t' (counter to fact) to consist only of the background information, the Q -function would not be the appropriate probability function to use. The historical fact that Smith had used the Q -function at t when Smith's 'total' evidence was the background information at t' would be irrelevant.

The same observation can be made from a slightly different point of view. Suppose that Jones does admit that the Q-function would be appropriate were the background information at t' the 'total' evidence at t' . Smith's argument would then succeed as a justification. But its success would not depend upon Jones admitting that at some time t^* prior to t' he had as 'total' evidence the background information at t' and had subsequently acquired e as the only new relevant evidence. Jones may never have been in such a position and may never have used the Q-function to determine fair betting quotients.

Thus, when conditionalising arguments of type (4) are used to justify probability judgments, temporal considerations are not of central relevance. The master function Q need not represent probability functions used by anyone at any time. This does not mean that they must be Carnapian functions which represent probabilities which should be assigned relative to no empirical evidence. They can represent probability assignments which should be made relative to the 'background information'—i.e. relative to all of the available evidence at the time when probability determinations are being made which is not explicitly taken into consideration at that time.

When viewed from this point of view, conditionalising arguments acquire critical importance in precisely those situations where Suppes claims they have none—namely in justifying probability judgments made at a given time. Where there is some doubt as to which of alternative systems of coherent probability judgments to use, attempts to justify adoption of one system will involve an explicit appeal to evidence e . As Suppes correctly points out, it would be absurd to require that the explicitly cited e be taken as the 'total' evidence, in some quite literal sense, available to the agents involved. The evidence not explicitly mentioned can be taken care of, however, by a master probability function Q representing probability judgments that would be made were the total evidence exactly the same save for the admission of e . Thus, conditionalising arguments can be used where appeals to coherence requirements alone will obviously be of little avail.

Moreover, the use of conditionalising arguments to justify probabilities assigned at one time by an appeal to probabilities assigned at another becomes a relatively unimportant application; for, as we have already seen, the success of the justification does not depend in any critical way upon whether the probabilities used at the earlier time were actually used then or not. If this is so, how can conditionalisation understood as argument to justify probability assignments be construed as a normative principle of learning?

There is an explanation for the confusion involved here which links it up with certain classical issues in epistemology. Just as Descartes attempted to follow a programme for justifying all beliefs held by a person at any time, Carnap and other writers on probabilistic inference seem to have understood a rational agent to be one whose every probability judgment and decision is at every time in his career as rational agent a justified one. At the inception of his career as rational agent, he adopts some justified probability function and revises his judgments over time in response to the acquisition of new evidence in a manner which will insure that at every later time his probability judgments will be justified. This means that an ideally rational agent will also be a rational learner. Since, for Carnap and many others, conditionalisation is the principle which

leads to justified revision of probabilities upon the acquisition of new evidence, it also becomes a principle of rational learning.

This conflation of justification and rational learning depends, however, on one crucial assumption, namely that at the inception of his career as rational agent the probability function used was a justified one. The Carnapian confirmation theory was intended to indicate how this condition could be met by characterising a probability function which would be justified were there no empirical evidence. Then even if the rational agent began his career with some empirical evidence, the justified function could itself be characterised by conditionalisation where the master function Q was the appropriate Carnapian confirmation function.

Advocates of a subjectivist interpretation of probability have been noted for their scepticism regarding the possibility of identifying a suitable Carnapian confirmation function. But rejection of such a special function should lead to an abandonment of the ideal of rational man which conflates rational inference with rational learning. If two agents with the same evidence can differ in their probability judgments (as long as they are coherent) in a way which prevents one system of judgments being justified over the other, then there seem to be no grounds for supposing that probability judgments made at a later time are justified only if they are obtained *via* conditionalisation from judgments made at an earlier time. What is allowable in the way of disagreements between Jones and Smith at the same time would appear to be permissible when the disagreement is between Jones at t and Jones at t' .

In spite of this, subjectivists have often treated intrapersonal disagreements over time differently from interpersonal differences at the same or different times. Such discrimination has been important in the defence of subjectivism against complaints that it is severely subjectivistic. As De Finetti and Savage have observed, if Jones and Smith disagree at time t , under certain kinds of conditions acquisition of new evidence can lead to agreement at time t' —provided that both Jones and Smith conditionalise relative to the new evidence and the probability functions they used at t . Thus, the subjectivists have continued to conflate the problem of justifying revisions in probability judgments with rational learning from experience even after they have abandoned the key element in the Carnapian programme which renders this identification plausible.

Suppes has apparently pushed this approach to its logical conclusion. If rational inductive inference is rational inductive learning, then an applicable theory of rational inference had better be based upon a thoroughgoing understanding of human abilities to process information and adjust their behaviour to it. Justification has been reduced to learning and the Carnapian variant of the Cartesian programme has been turned on its head. In effect, the problem of providing criteria for justifying the selection of one coherent system of probability judgments from many has been dropped.

Critics of the subjectivist approach have often located the source of trouble in the interpretation of probability as an expression of a fair betting quotient and in the Bayesian approach to decision theory. However, it seems to me that the dubious implications of Suppes's approach can be avoided without attacking the subjectivist interpretation *per se*. One can, instead, question the Carnapian conception of justification and rationality.

The debatable assumption of this approach is the requirement that all of the beliefs, actions and probability judgments adopted by a rational agent at any time be justified. This assumption implies that the premisses of any successful justification must themselves be justified. If infinite regress and vicious circularity are to be avoided, some assumptions must be self justifying. The troubles which analogous views in classical epistemology have engendered are notorious and need not be reviewed here.

As Charles Peirce emphasised a long time ago, an alternative conception of rationality and justification is possible. The demand for justification of some judgment or decision at a given time t is successfully met by an argument provided that the premisses of the argument and the principles by means of which the conclusion is obtained are not at that time in need of justification—i.e. provided that there is no serious difficulty or problem raised at that time regarding their acceptability. Adoption of this approach is compatible—indeed requires admitting—that premisses which do not stand in need of justification may be subjected to critical scrutiny at some later time. But the crucial point is that a rational agent is no longer one whose every belief at every time is a justified or even a justifiable one. Instead, he may be viewed as a person who adjusts his beliefs and decisions to the requirements of justified belief and action when justification is legitimately demanded.

Thus, conditionalising arguments can be used to justify choosing one coherent system of probability judgments at a given time t rather than some other set given a noncontroversial master function expressing a judgment regarding the probabilities which should be assigned were the background evidence the 'total' evidence not subject to serious question and were the evidence e all of the explicitly identified evidence not in the background. Success of an argument of this sort as a Peircean justification does not depend upon the master Q -function being used to determine fair betting quotients at some earlier time. The application of conditionalisation in such a context cannot be construed as a prescription for ideal learning.

Of course, in defending the applicability of conditionalisation or any other principle of rational inference, psychological considerations do become relevant. Agents must be capable of using the principles to determine which conclusions are justifiable and adjusting their opinions and decisions accordingly. But since theories of rational inference are not theories of rational learning, learning theory is not central to a study of rational inference in the manner in which Suppes suggests it is. Detailed consideration of the ways in which psychological considerations can enter into discussions of rational inference is a difficult matter and space will not allow a serious discussion here. But enough has been said to suggest that students of probabilistic inference need not rush to a study of learning theory just yet.

The chief point of this discussion has been to suggest that advocates of a Bayesian approach to decision theory and subjectivist interpretations of probability need not be committed to the form of psychologism which Suppes advocates. It has not been intended as a blanket defence of Bayesianism in general or conditionalisation in particular. But it is worth noting that conditionalisation does have its most effective application precisely in cases where statistical syllogisms are employed.

Suppose that Jones and Smith both agree to admit the statistical hypothesis $p(A, B) = r$ as evidence. Under a wide variety of conditions in which, among other things, the background information available to Jones and Smith may vary, they may nonetheless agree that were the background information and the statistical hypothesis the 'total' evidence, they would assign a conditional probability to the hypothesis that a is an A given that a is a B equal to r . Together with the assumption that a is a B and that no evidence other than the background and statistical hypothesis is relevant, conditionalisation will justify assigning a probability of r to the hypothesis that a is an A.

I do not mean to gloss over difficulties in stating the conditions of applicability of statistical syllogistic arguments. Precise specification of these conditions is no easy matter as the recent efforts of Kyburg and Hempel testify. But it does seem clear that such arguments do succeed on many occasions in providing justifications of probability assignments used to determine fair betting quotients. Indeed, they provide the best illustrations of the use of conditionalisation as a principle of justification. Their use in statistical prediction and explanation suffices in itself to cast doubt on Suppes's critique of the statistical syllogism.

The Peircean conception of justification just outlined presupposes that both the background information and explicitly identified evidence consists of propositions implicitly or explicitly accepted as true without serious question, i.e. an evidence sentence is one which is accepted as true and, in addition, one whose truth does not stand in need of further testing. However, because what is exempt from serious difficulty at one time may be questioned later on, a theory of rational inference must not only determine how probabilities are to be assigned to hypotheses on given evidence but when sentences are justifiably admitted into evidence on other evidence which does not entail them. This leads to consideration of 'inductive acceptance rules' which Bayesian theory itself is impotent to handle.

Jaakko Hintikka together with his associates Risto Hilpinen and Juhani Pietarinen explore various problems connected with inductive acceptance rules in two interesting papers in this volume. The paper by Hintikka and Hilpinen ('Knowledge, Acceptance and Inductive Logic') considers the apparent conflict between the requirement of deductive closure (which demands that the deductive consequences of a set of accepted sentences also be accepted) and the multiplication theorem for probabilities. Hintikka and Pietarinen ('Semantic Information and Inductive Logic') exploit decision theoretic ideas together with the concept of epistemic utility to offer an account of the way in which the demand for information controls criteria for rational inference. Both papers rely heavily on Hintikka's work on confirmation measures which is summarised and generalised in another paper by Hintikka ('A Two Dimensional Continuum of Inductive Methods') and extended in various ways in two papers by Hilpinen and Raimo Tuomela.

In addition to the papers which have been mentioned, there is a discussion of the simplicity of state descriptions by Kurt Walk, a formal examination of the properties of probabilistic measures of 'evidential strength' as compared to informational measures by Hakan Törnebohm, three interesting papers by Max Black, Patrick Suppes and G. H. von Wright on the paradox of confirmation, an important paper by Dana Scott and Peter Krauss concerning the application of probability measures to formulae of suitably formalised logical calculi, and an

effort to avoid the paradoxes of material implication with the aid of probabilities by Ernest Adams.

The editors of this volume are to be congratulated on the remarkably high quality of the majority of contributions. All too often writers on probability and induction find it difficult to avoid engaging in philosophically pointless displays of technical virtuosity. There is plenty of virtuosity in this book. But with few exceptions, a philosophical issue of importance controls the technicalities. This volume deserves the serious attention of all students of probability and induction.

ISAAC LEVI

Western Reserve University,
Case Institute of Technology

Gordon J. DiRenzo (ed.): *Concepts, Theory, and Explanation in the Behavioral Sciences* (New York: Random House, 1966, \$4.95). Pp. xii + 302.

The book is another of those reports on symposia which seem to emerge from the U.S.A. with increasing frequency. The theme of the present symposium is 'Conceptual Definition in the Behavioural Sciences', although behind this lies a more general concern with theory construction in this field. The participants in this particular symposium are mainly sociologists although there are also included some social psychologists and one philosopher—Alan Ross Anderson. There are contributions from individuals, or in some cases groups of individuals, and the editor has prefixed each contribution with a summary of what is said, together with a certain amount of discussion aimed at putting the contribution in its place in the general scheme. There is also a record of a brief discussion, which is of doubtful value, and a final lengthy contribution from the editor in which he attempts to sum up the issues. As is usual in this kind of situation the individual contributors tend to speak in the main about what particularly interests them. Thus Charles Ackerman and Talcott Parsons speak about the notion of a 'social system' on the lines of the latter's 'action theory'; Leslie A. White, the anthropologist and 'culturologist', gets very involved in what seem to me largely imaginary difficulties over the term 'culture', while a group of psychologists, Levinson, Sharaf and Gilbert, describe the development and applications of the concept of 'intraception'—a concept derived from Jung's 'introversion'. Levinson, incidentally, employs the notion of 'intraception' *ad nauseam* in the discussion, in order to pigeon-hole and perhaps provide an assessment of the contributions of his fellow symposiasts. The remaining contributions—from the psychologist Muzafer Sherif, from Alan Ross Anderson and a sociologist co-contributor, Omar K. Moore, from a sociologist, Paul F. Lazarsfeld, and from the editor himself, who is likewise a sociologist—are more explicitly methodological in aim.

There is no doubt that the editor works very hard in order to abstract from all this some general conclusions about theory construction in the field, and his own final contribution has a great deal of interest from this point of view. It has, however, some strange features. Thus there is a preoccupation with a supposed

distinction between real and nominal definitions and in consequence between 'real and nominal concepts'. Real concepts, we are told (p. 268), are necessarily true; they refer to ontological reality. These are not (p. 274) 'given at the outset of inquiry; they result only from empirical investigation of the phenomena in question' and the 'validity of a theoretical system, and of its conceptual apparatus, is obtained by empirical confirmation' (p. 276)—something which is then identified with prediction. There are many questions begged in all this, and I suspect that the discussion shows a fundamental lack of clarity about what is involved in a scientific theory. The preoccupations with operationalism, which beset psychology some twenty years ago, are to be met with here, too, and in Lazarsfeld's contribution, although Lazarsfeld eventually emerges with the claim that concepts in the behavioural sciences can never be defined precisely but only in probabilistic terms.

What is, however, lacking in general, unless this is to be found in the more opaque pronouncements of Parsons etc., is a systematic consideration of whether the subject-matter of the social sciences presents in itself problems for its methodology. There is no consideration, for example, of the kinds of question raised by P. G. Winch in his *The Idea of a Social Science*. Indeed there is in DiRenzo's concluding paragraphs a specific repudiation of any suggestion that the methodology of the behavioural sciences should not be scientific. But this again begs questions, since it does not follow from the fact that scientific methods can be used in sociology that the theories which can be constructed in that field may not have peculiarities all of their own. Thus it is not enough to say, as is said on page 259, 'After one has gone about deciding what constitutes explanation, and how best to achieve it, then the next logical problem that remains to be settled is the question of what it is that is to be explained'. This may well be a serious case of putting the cart before the horse.

Anderson comments in his contribution to the discussion that the papers involve much traditional philosophy and comparatively little philosophy of science in the modern sense; there is little reference to philosophers like Popper, who are leading figures in the philosophy of science. I think that this is generally true, although DiRenzo's final contribution does something to redeem the situation. Nevertheless it must be said that, *qua* philosophy of science, the discussion is relatively crude. One of the more interesting points, which is an exception to this general rule, is Anderson and Moore's claim about the value of mechanical models in the explanation of human behaviour. They say (p. 87): 'The virtue of attempts at mechanical simulation is that they *may*, with luck, suggest the kind of theoretical problems which, if solved, would explain both the original behaviour and that of the mechanical analogue. But theories, or explanations, make *assertions*; mechanical models do not.' This seems to me to be correct and it is as well that it should be said and resaid.

Apart, however, from a small number of incidental points of this kind, I do not think that the symposium as a whole contributes much illumination. The participants are too concerned either with rather local issues arising from their own work or with methodological issues which are too general and too crudely formulated to fit in with the real problems which face the philosopher of social science. It is perhaps a pity that more philosophers were not present, since the issues discussed are undoubtedly philosophical and sociologists are not exactly

renowned for the clarity of their discussions of their own subject. In sum, the essays do little to fulfil the hope of the editor that they will 'contribute to a clarification that may be of some significant benefit, not only to the behavioural sciences, but to all the other areas of scientific inquiry as well'.

D. W. HAMLYN
Birkbeck College,
University of London

Paul F. Schmidt: *Perception and Cosmology in Whitehead's Philosophy* (Rutgers University Press, 1967, \$9.00). Pp. xvi+192.

Professor Schmidt has written a book which can be of interest to those concerned with the exegesis of Whitehead's metaphysics and who have already worked themselves into the system and are prepared to argue its finer points in Whitehead's own terminology. What it does not do is to try to give an estimate from an independent viewpoint. This need not be an unsympathetic viewpoint, but one which stands back from the system and tries to discuss the meaning of some of Whitehead's key terms, e.g. 'feeling', in language other than his own. Professor Schmidt does not think this is possible because 'ordinary philosophical language' 'is based on a subject-predicate type of statement which can never hope to express a process philosophy' (p. 128). Will this do? After all, Whitehead himself writes in subject-predicate statements: 'actual entities enjoy feelings', 'are in community with each other' and so on; he does not write in the style of James Joyce or Gertrude Stein. Professor Schmidt seems to be assuming that the grammar in which we talk is mirrored in how we see what we are talking about. Isn't it possible to use subject-predicate statements to talk about processes in interrelation, as other process philosophers besides Whitehead have done?

This book will therefore hardly appeal to philosophers of science who are not already fairly closely involved in the exegesis of Whitehead. Those who are will find that Schmidt does indeed show how some critics have taken insufficient note of what Whitehead has said in some part of his extensive writings, which Schmidt obviously knows inside out. They will also find references in the footnotes and bibliography to a large number of articles on Whitehead not all of which may have come their way. But the failure to discuss closely, as distinct from expounding, means that Schmidt seems unaware of difficulties others have found in notions such as 'mutual immanence' and 'objective immortality'. This is a pity, for if we could get their meaning out, they might well be found to have an importance beyond the world of Whitehead's own system.

DOROTHY EMMET
Cambridge

Abstracts

Philosophy of Science, 34 (1967), No. 2

Paul E. Meehl, 'Theory-testing in psychology and physics: a methodological paradox'

Because physical theories typically predict numerical values, an improvement in experimental precision reduces the tolerance range and hence increases corroborability. In most psychological research, improved power of a statistical design leads to a prior probability approaching $\frac{1}{2}$ of finding a significant difference in the theoretically predicted direction. Hence the corroboration yielded by 'success' is very weak, and becomes weaker with increased precision. 'Statistical significance' plays a logical role in psychology precisely the reverse of its role in physics. This problem is worsened by premium placed on experimental 'cuteness' and a free reliance upon *ad hoc* explanations to avoid refutation.

Brian Ellis and Peter Bowman, 'Conventionality in distant simultaneity'

In his original paper of 1905, 'On the electrodynamics of moving bodies', Einstein described a procedure for synchronising distant clocks at rest in any inertial system K. Clocks thus synchronised may be said to be in *standard signal synchrony* in K. It has often been claimed that there are no logical or physical reasons for preferring standard signal synchronisations to any of a range of possible non-standard ones. In this paper, the range of consistent non-standard signal synchronisations, first for any one inertial system, and second for any set of such systems, is investigated, and it is shown that the requirement of consistency leaves much less room for choice than is commonly supposed. Nevertheless consistent non-standard signal synchronisations appear to be possible. However, it is also shown that good physical reasons for preferring standard signal synchronisations exist, if the Special Theory of Relativity yields correct predictions.

The thesis of the conventionality of distant simultaneity espoused particularly by Reichenbach and Grünbaum is thus either trivialised or refuted.

Kenneth F. Schaffner, 'Approaches to reduction'

Four current accounts of theory reduction are presented, first informally and then formally: (1) an account of direct theory reduction that is based on the contributions of Nagel, Woodger, and Quine, (2) an indirect reduction paradigm due to Kemeny and Oppenheim, (3) an 'isomorphic model' schema traceable to Suppes, and (4) a theory of reduction that is based on the work of Popper, Feyerabend, and Kuhn. Reference is made, in an attempt to choose between these schemas, to the explanation of physical optics by Maxwell's electromagnetic theory, and to the revisions of genetics necessitated by partial biochemical reductions of genetics. A more general reduction schema is proposed which: (1) yields as special cases the four reduction paradigms considered above, (2) seems to be in better accord with both the canons of logic and actual scientific practice, and (3) clarifies the problems of meaning variance and ontological reduction.

Bernard Suits, 'What is a game?'

By means of a critical examination of a number of theses as to the nature of game-playing, the following definition is advanced: 'To play a game is to engage in activity directed toward bringing about a specific state of affairs, using only means permitted by specific rules, where the means permitted by the rules are more limited in scope than they would be in the absence of the rules, and where the sole reason for accepting such limitation is to make possible such activity.'

Hugh Lehman, 'Are biological species real?'

Difficulties with the typological concept of species led biologists to reject the 'typological' presupposition of an archetype which is manifest in each member of a species. The resulting concept of species, which is here called the phenotypic species concept, is considered as implying that biological species are not real. Modern population thinking has given rise to the concept of a species as a gene-pool. This modern concept is contrasted here with the phenotypic concept in light of some general criteria for evaluating species concepts and is shown to be more satisfactory. Finally, it is held that to ask if a species is real is to ask whether the species grouping arrived at by applying the principles involved in the species concept corresponds with groups of organisms amongst which important biological relationships exist. It is argued that in this sense species, as defined by the gene-pool concept, are certainly real.

Philip K. Bock, 'Three descriptive models of social structure'

In this article, three descriptive models are described and compared in an attempt to discover the minimum number of concepts and relations necessary to an abstract description of social structure. It is argued that a 'situational model' which uses the concepts of social time, social space and social role together with the relations of internal structure and external distribution is both adequate to the task and superior to the two other models. The advantages and implications of the situational model are discussed.

Philosophy of Science, 34 (1967), No. 3

John Earman, 'On going backward in time'

This paper presents a critical examination of claims advanced by several philosophers to the effect that 'time travel' represents a physical possibility and that the interpretation of certain actually observed phenomena in terms of 'time travel' is both legitimate and advantageous. It is argued that (a) no convincing motivation for the introduction of the time travel hypothesis has been presented; (b) no coherent and interesting sense of 'going backward in time' has been supplied which makes 'time travel' compatible with Special Relativity; (c) even the conceptual possibility of 'time travel' is an unsettled and somewhat nebulous question.

I. C. Jarvie, 'On theories of fieldwork and the scientific character of social anthropology'

The following intellectual as opposed to practical reasons for all anthropologists doing fieldwork are examined: fieldwork (1) records dying societies, (2) corrects ethnocentric bias, (3) helps put customs in their true context, (4) helps get the 'feel' of a place, (5) helps to get to understand a society from the inside, (6) enables appreciation of what translating one culture into terms of another involves, (7) makes one a changed man, (8) provides the observational, factual basis for generalisations. None of these is found sufficient to make fieldwork imperative for all anthropologists, although they are quite sufficient to allow that it is imperative for anthropology as a whole that fieldwork in some form by some people continues. In place of the view of fieldwork as an essential preparation for doing anthropology, an alternative role for it is explored: namely as a testing procedure. The implications of this—that the study of problems and the articulation of theories can usefully proceed prior to or even independently of fieldwork—are drawn out, and a new institution of selective fieldwork is proposed.

Vivian Charles Walsh, 'On the significance of choice sets with incompatibilities'

The axiom of comparability has been a fundamental part of mathematical choice theory from its beginnings. This axiom was a natural first assumption for a theory of choice originally constructed to explain decision making where other assumptions such as continuous divisibility of choice spaces could legitimately also be made. Once the generality of application of formal choice theory becomes apparent, it also becomes apparent that

both continuity assumptions and the axiom of comparability may be unduly restrictive and lead to the neglect of decision situations which are important and which can be handled on a modified axiom set. These considerations bear on the philosophical analysis of the concept of rational decision.

Laird Addis, 'Particulars and acquaintance'

Philosophers who hold that the correct ontological analysis of things includes both properties and particulars have often been pressed to 'show' the particular. If we are not *acquainted* with them, it is argued, then we should not suppose that they exist. I argue that, while we do have good and sufficient reasons for supposing there to be particulars, we are not acquainted with them. To suppose that we are acquainted with them is to treat particulars as if they were properties and to fail to realize how radically different particulars are from properties. The relevance of these matters to some considerations of 'simplicity' and the principles of empiricism is explored.

Edwin B. Allaire, 'Things, relations and identity'

Philosophers have long believed that if the Principle of the Identity of Indiscernibles were *logically* true, there would be no problem of individuation. I show (a) that if spatial relations are, as seems plausible, of such a *nature* that it makes no sense to say of one thing that it is related to itself, then the Principle is a *logical* truth, asserting that a certain kind of state of affairs is impossible because the kind of sentence purporting to express it is ill-formed, and (b) that even if the Principle were such a truth the problem of individuation would remain. I then examine some of what Leibniz and Wittgenstein have said about the notions of individuation, difference and non-identity.

Recent Publications

RECENT PUBLICATIONS ON THE PHILOSOPHY OF SCIENCE

BOOKS RECEIVED FOR REVIEW

- Alice Ambrose: *Essays in Analysis*. (London: G. Allen and Unwin; New York: Humanities Press. 1966. 40s.) pp. 262
- John P. Anton, ed: *Naturalism and Historical Understanding*. (State University of New York Press. 1967. \$10.00) pp. x+323
- St Thomas Aquinas: *Summa Theologiae*, vol. 10, *Cosmogony*. Translated William A. Wallace. (London and New York: Eyre and Spottiswoode and McGraw-Hill. 1967. 42s.) pp. xxiii+255
- Ian G. Barbour: *Issues in Science and Religion*. (London: S.C.M. Press. 1966. 45s.) pp. x+470
- Antonín Bartonek: *Development of the Long-Vowel System in Ancient Greek Dialects*. (Praha: Státní Pedagogické Nakladatelství. 1966) pp. 199
- Saul A. Basri: *A Deductive Theory of Space and Time*. (Amsterdam: North Holland Publishing Co. 1966) pp. xi+163
- Leonardo Benevolo: *The Origins of Modern Town Planning*. (M.I.T. Press. 1967. \$5.50) pp. xi+154
- J. G. Bennett: *The Dramatic Universe*. (London: Hodder and Stoughton. 1966. 70s.) pp. xxxvi+315
- J. G. Bennett: *The Dramatic Universe*. (London: Hodder and Stoughton. 1967. 84s.) pp. i+462
- Gustav Bergmann: *The Metaphysics of Logical Positivism*. (University of Wisconsin Press. 1967. \$2.95) pp. xii+340
- Gustav Bergmann: *Realism: A Critique of Brentano and Meinong*. (University of Wisconsin Press. 1967. \$3.75) pp. vii+458
- J. D. Bernal: *The Social Function of Science*. (M.I.T. Press. 1967. \$3.95) pp. xxxvi+482
- W. H. Brock, ed.: *The Atomic Debates*. (Leicester University Press. 1967. 35s.) pp. ix+186
- Justus Buchler: *Metaphysics of Natural Complexes*. (New York and London, Columbia University Press. 1966. 45s.) pp. viii+195
- Mario Bunge, ed: *Delaware Seminar in the Foundations of Physics*. (Berlin and New York: Springer-Verlag. 1967. DM38; \$9.50) pp. xii+193
- Mario Bunge: *Scientific Research I. The Search for System*. (Berlin and New York: Springer-Verlag. 1967. DM84; \$21.00) pp. xii+536
- Mario Bunge: *Scientific Research II. The Search for Truth*. (Berlin and New York: Springer-Verlag. 1967. DM68; \$17.00) pp. viii+374
- Rudolf Carnap: *Logical Structure of the World, and Pseudoproblems in Philosophy*. (London: Routledge and Kegan Paul. 1967. 60s.) pp. xxvi+364
- Hector-Neri Castaneda: *Intentionality, Minds and Perception*. (Detroit: Wayne State University Press. 1967. \$7.95) pp. 402
- Rollin Chambliss: *Meaning for Man*. (New York: Philosophical Library. 1966. \$4.50) pp. xi+191
- Robert G. Colodny, ed: *Mind and Cosmos*. (University of Pittsburgh Press. 1966. \$8.00) pp. xviii+362
- J. B. Conant: *Scientific Principles and Moral Conduct*. (Cambridge University Press. 1967. 6s.) pp. 48
- Irving M. Copi and Robert W. Beard, eds: *Essays on Wittgenstein's Tractatus*. (London: Routledge and Kegan Paul. 1966. 50s.) pp. x+414
- James W. Cornman: *Metaphysics, Reference and Language*. (Yale University Press. 1966. 56s.) pp. xxi+288

- Descartes: *The Philosophical Works of Descartes*. 2 vols. Translated by Elizabeth S. Haldane and G. R. T. Ross. (Cambridge University Press. 1967. 12s. 6d. each vol.) pp. vi+452; 380
- Harald Dickson: *Variable, Function, Derivative: A Semantic Study in Mathematics and Economics*. (Handelshögskolan i Göteborg, Skrifter 1967: 1, SKr.30) pp. 176
- Gordon J. Dizenzo: *Concepts, Theory and Explanation in the Behavioral Sciences*. (New York: Alfred A. Knopf. 1966. \$4.95) pp. xii+302
- Lloyd D. Easton: *Hegel's First American Followers, the Ohio Hegelians*. (Ohio University Press. 1967. \$7.00) pp. ix+353
- Walter M. Elsasser: *Atom and Organism*. (Princeton University Press. 1966. 36s. \$4.50) pp. ix+143
- J. K. Feibelman: *Moral Strategy*. (The Hague: Martinus Nijhoff. 1967. Gs.37.25) pp. xi+325
- Paul K. Feyerabend and Grover Maxwell, eds: *Mind, Matter and Method: Essays in Philosophy and Science in Honor of Herbert Feigl*. (Minnesota University Press, Oxford University Press. 1967. 78s.) pp. 524
- Richard Feynman: *The Character of Physical Law*. (M.I.T. Press. 1967. \$2.45) pp. 173
- Gottlob Frege: *The Basic Laws of Arithmetic*. Translated and Edited Montgomery Furth. (Cambridge University Press, California University Press. 1967. 16s.) pp. lxiii+144
- I. M. Gelfand, E. G. Glagoleva and A. A. Kirillov: *The Method of Coordinates*. Translated L. Cohn and D. Sookne. (M.I.T. Press. 1967. \$1.50) pp. viii+69
- C. C. Gillespie: *The Edge of Objectivity: An Essay in the History of Scientific Ideas*. (Princeton University Press, Oxford University Press. 1967. 22s. 6d.) pp. 562
- Bentley Glass: *Science and Ethical Values*. (Oxford University Press. 1966. 28s.) pp. ix+101
- Alvin W. Gouldner: *Enter Plato: Classical Greece and the Origins of Social Theory*. (London: Routledge and Kegan Paul. 1967. 55s.) pp. ix+407
- G. R. Grice: *The Grounds of Moral Judgment*. (Cambridge University Press. 1967. 37s. 6d.) pp. ix+208
- D. ter Haar: *The Old Quantum Theory*. (London: Pergamon Press. 1967. 21s.) pp. x+206
- Otto Hahn: *A Scientific Autobiography*. (London: MacGibbon and Kee. 1967. 45s.) pp. xxiv+296
- G. H. Hardy: *A Mathematician's Apology*. (Cambridge University Press. 1967. 15s.) pp. 153
- Michael M. Hare: *Microcosm and Macrocosm*. (New York: The Julian Press. 1966. \$15.00) pp. xxv+291
- Errol E. Harris: *The Foundations of Metaphysics in Science*. (London: G. Allen and Unwin. 1965. 63s.) pp. 493
- Robert S. Hartman: *The Structure of Value: Foundations of Scientific Axiology*. (Southern Illinois University Press. 1967. \$10.00) pp. xix+384
- Jean van Heijenoort: *From Frege to Gödel*. (Harvard University Press, Oxford University Press. 1967. £7 8s.) pp. x+660
- Barnabas Hughes, ed: *Regiomontanus on Triangles*. (University of Wisconsin Press. 1967. \$10.00) pp. vii+298
- Stanley L. Jaki: *The Relevance of Physics*. (Chicago and London: University of Chicago Press. 1966. \$12.50) pp. 604
- Richard C. Jeffrey: *The Logic of Decision*. (New York: McGraw-Hill. 1965. \$7.95) pp. xi+201
- P. Kaim-Caudle: *Social Policy in the Irish Republic*. (London and New York: Routledge and Kegan Paul and the Humanities Press. 1967. 15s.) pp. 120
- Kant: *La Forma e i Principi del Mondo Sensibile e del Mondo Intelligibile*. Traduzione Ada Lamacchia. (Padova: Liviana Editrice. 1967. L.1000) pp. vii+119
- Robert Kar: *Einsteins Relativität eine Energiebedingte Illusion*. (Leiden: A. W. Sijthoff. 1965. Gs.10) pp. 159
- C. W. Kilminster: *Sir Arthur Eddington*. (Oxford and London: Pergamon Press. 1966. 21s.) pp. vi+279
- Rin'ichi Kittaka: *New Method of Philosophy*. (Tokyo: Risosha. 1967.) pp. 116
- J. Klir and M. Valach: *Cybernetic Modelling*. (London: Iliffe. 1967. 63s.) pp. 437
- David M. Knight: *Atoms and Elements*. (London: Hutchinson. 1967. 30s.) pp. 167

- Alexandre Koyré: *Dal Mondo del Pressappoco all'Universo della Precisione*. (Torino: Giulio Einaudi, 1967. L.500) pp. 134
- P. Kratochvíl: *Crystals*. (London and Prague: Iliffe and SNTL, 1967. 15s.) pp. 112
- G. Kreisel and J. L. Krivine: *Éléments de Logique Mathématique*. (Paris: Dunod, 1967. F58) pp. viii+212
- Franz von Kutschera: *Elementare Logik*. (Vienna and New York: Springer-Verlag, 1967. \$12.00) pp. viii+392
- J. Kvasnica: *Physical Fields*. (London: Iliffe, 1966. 17s. 6d.) pp. 183
- J. Kvasnica: *Structure of the Atomic Nucleus*. (London and Prague: Iliffe and SNTL, 1967. 15s.) pp. 121
- Grace A. de Laguna: *On Existence and the Human World*. (Yale University Press, 1966. 48s., \$6.50) pp. xi+267
- Imre Lakatos, ed: *Problems in the Philosophy of Mathematics*. (Amsterdam: North-Holland Publishing Co. 1967. 72s.) pp. xv+241
- Hugues Leblanc: *Techniques of Deductive Inference*. (London: Prentice-Hall International, 1966. 48s.) pp. vii+216
- Martin Levey, ed: *Archaeological Chemistry*. (Philadelphia: University of Pennsylvania Press, 1967. \$8.50) pp. 365
- Isaac Levi: *Gambling with Truth: An Essay on Induction and the Aims of Science*. (New York and London: Alfred A. Knopf and Routledge and Kegan Paul, 1967. 63s.) pp. xiv+246+v.
- Salomon E. Liverhant: *Outline of Atomic Physics*. (London: Iliffe, 1966. 30s.) pp. 315
- M. Stanley Livingston: *The Development of High-Energy Accelerators*. (New York: Dover, 1966. 20s.) pp. xi+317
- Raphael Lowe, ed: *Studies in Rationalism, Judaism and Universalism. In memory of Leon Roth*. (London and New York: Routledge and Kegan Paul, 1966. 50s.) pp. xiii+357
- A. R. Louch: *Explanation and Human Action*. (Oxford: Blackwell, 1967. 37s. 6d.) pp. viii+243
- Storrs McCall, ed: *Polish Logic 1920-1939*. (Oxford University Press, 1967. 90s.) pp. vii+406
- Alasdair MacIntyre: *A Short History of Ethics*. (London: Routledge and Kegan Paul, 1966. 30s.) pp. viii+280
- Charles McCreery: *Science, Philosophy and ESP*. (London: Faber and Faber, 1967. 32s. 6d.) pp. 199
- Albertus Magnus: *Book of Minerals*. Translated Dorothy Wyckoff. (Oxford University Press, 1967. 84s.) pp. xiii+309
- Maurice Mandelbaum: *Philosophy, Science, and Sense Perception*. (Baltimore and London: Johns Hopkins Press and Oxford University Press, 1967. 20s.) pp. x+262
- D. H. Martin: *Magnetism in Solids*. (London: Iliffe 1967. 120s.) pp. 452
- Vittorio Mathieu, ed: *Spazio e Cielo*. (Firenze: Sansoni, 1966. L.2800) pp. 200
- Anton Meinrad Meier: *Das Peccatum Mortale ex Toto Genere Suo*. (Regensburg: Verlag Friedrich Puster, 1966. DM45) pp. 405
- A. Michalowicz: *Kinematics of Nuclear Reactions*. Edited R. A. Giles. (London: Iliffe, 1967. 30s.) pp. 140.
- Andrzej Mostowski: *Thirty Years of Foundational Studies*. (Oxford: Blackwell, 1966. 37s. 6d.) pp. 180
- J. Neyman and E. S. Pearson: *Joint Statistical Papers*. (Cambridge University Press, 1967. 45s.) pp. 299
- Marjorie Hope Nicholson: *Newton Demands the Muse*. (Princeton University Press, Oxford University Press, 1967. \$2.95, 22s. 6d.) pp. xi+178
- H. A. Nielsen: *Methods of Natural Science*. (London: Prentice-Hall International, 1967. 14s.) pp. xv+70
- Jane M. Oppenheimer: *Essays in the History of Embryology and Biology*. (M.I.T. Press, 1967. \$12.50) pp. ix+374
- M. H. Pappworth: *Human Guinea Pigs* (London: Routledge and Kegan Paul, 1967. 32s.) pp. xi+228
- D. H. Parkinson and B. E. Mulhall: *The Generation of High Magnetic Fields*. (London: Iliffe, 1967. 70s.) pp. xii+165

- Karle Patek: *Lasers*. Edited V. W. Rampton. (London and Prague: Iliffe and SNTL. 1967. 45s.) pp. 288
- Pearson: *Selected Papers of E. S. Pearson*. (Cambridge University Press. 1966. 40s.) pp. vii+327
- J. A. Philip: *Pythagoras and Early Pythagoreanism*. (Oxford University Press. 1967. 35s.) pp. x+222
- Michael Polanyi: *The Tacit Dimension*. (London: Routledge and Kegan Paul. 1967. 16s.) pp. xi+108
- Margaret Purver: *The Royal Society: Concept and Creation*. (London: Routledge and Kegan Paul. 1967. 35s.) pp. xvii+246
- Robert N. Rapoport: *Community as Doctor: New Perspectives on a Therapeutic Community*. (London: Tavistock Publications. 1967. 25s.) pp. x+325
- Oliver L. Reiser: *Cosmic Humanism*. (Massachusetts: Schenkman Publishing Co. 1966. \$8.95) pp. xxvi+576
- Nicholas Rescher: *Distributive Justice*. (New York: Bobbs-Merrill. 1966. \$6.00) pp. xvi+166
- Nicholas Rescher: *The Philosophy of Leibniz*. (New Jersey: Prentice-Hall. 1967) pp. viii+168
- T. Rivers: *Reflections on a Life in Medicine and Science*. (M.I.T. Press. 1967. \$17.50) pp. xviii+682
- Nesta Roberts: *Mental Health and Mental Illness*. (London and New York: Routledge and Kegan Paul and the Humanities Press. 1967. 15s.) pp. vi+80
- P. Rogers and G. A. Stephens: *Worked Examples in Modern Physics*, 2 vols. (London: Iliffe. 1967. 18s. 6d. each vol.) pp. 16+227
- Nathan Rotenstreich: *On the Human Subject*. (Illinois. 1966.) pp. vii+226
- Dagobert D. Runes: *Treasury of Thought*. (New York: Philosophical Library. 1966. \$6.00) pp. 395
- Wesley C. Salmon: *The Foundations of Scientific Inference*. (University of Pittsburgh Press. 1966/67. \$2.25) pp. 157
- Jean-Paul Sartre: *Of Human Freedom*. Edited W. Baskin. (New York: Philosophical Library. 1966. \$4.75) pp. xiii+158
- Eike von Savigny: *Die Überprüfbarkeit der Strafrechtssätze*. (München: Verlag Karl Alber. DM14.80) pp. 151
- Richard Schlegel: *Completeness in Science*. (New York: Appleton-Century-Crofts. 1967) pp. xi+280.
- Robert E. Schofield, ed: *A Scientific Autobiography of Joseph Priestley, 1733-1804*. (M.I.T. Press. 1966. \$13.50) pp. xiv+415
- J. F. Scott, ed: *The Correspondence of Isaac Newton, Vol. IV, 1694-1709*. (Cambridge University Press. 1967. £11 11s.) pp. xxxii+578.
- G. L. S. Shackle: *The Years of High Theory*. (Cambridge University Press. 1967. 45s.) pp. vii+328
- Baruch Spinoza: *Letters to Friend and Foe*. Edited Dagobert D. Runes. (New York: Philosophical Library. 1966. \$3.75) pp. 109
- J. B. Stallo: *The Concepts and Theories of Modern Physics*. Edited Percy W. Bridgman. (Harvard University Press, Oxford University Press. 1966. 16s.) pp. xxix+325
- Robert Sternfeld: *Frege's Logical Theory*. (Southern Illinois University Press. 1966. \$8.50) pp. xiii+200
- Julius A. Stratton: *Science and the Educated Man*. (M.I.T. Press. 1966. 40s.) pp. viii+186
- A. M. Taylor: *Imagination and the Growth of Science*. (London: John Murray. 1966. 16s.) pp. 110
- Merle B. Turner: *Philosophy and the Science of Behavior* (New York: Appleton-Century-Crofts. 1967. \$9.50) pp. xvii+539
- B. Ureosik: *Physical Electronics*. (London and Prague: Iliffe and SNTL. 1967. 17s. 6d.) pp. 189
- James R. Ware, trans: *Alchemy, Medicine, Religion in the China of A.D. 320*. (M.I.T. Press. 1966. \$15.00) pp. xiv+388.
- Paul Weingartner: *Beschreibung, Analytizität und Existenz*. (Salzburg: Anton Pustet. 1966. DM58) pp. 416

92 *Recent Publications*

- Alan R. White: *The Philosophy of Mind*. (New York: Random House. 1967. \$1.95) pp. viii+178
- G. J. Whitrow, ed: *Einstein: The Man and His Achievement*. (B.B.C. 1967. 18s. 6d.) pp. xv+94
- Alvin M. Weinberg: *Reflections on Big Science*. (M.I.T. Press. 1967. \$5.95) pp. ix+182
- A. J. Willcocks: *The Creation of the National Health Service*. (London and New York: Routledge and Kegan Paul and the Humanities Press. 1967. 15s.) pp. vii+118
- Ludwig Wittgenstein: *Bemerkungen über die Grundlagen der Mathematik (Remarks on the foundations of mathematics)*. Edited G. H. von Wright and others. Translated G. E. M. Anscombe. (M.I.T. Press. 1967. \$3.45) pp. xix+204
- Ludwig Wittgenstein: *Zettel*. Edited G. E. M. Anscombe and G. H. von Wright. (Oxford: Basil Blackwell. 1967. 37s. 6d.) pp. v+124
- Herman A. O. Wold and others: *Model Building in the Human Sciences*. (Monaco: Editions 'Sciences Humaines'. 1966.) pp. xii+321
- W. D. Wright: *The Rays are not Coloured*. (London: Hilger and Watts. 1967. 35s.) pp. x+154

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

Honorary Member

At its Annual General Meeting on 18 March 1968, Honorary Membership was conferred on Professor Herbert Dingle in recognition of his outstanding services to the Society, especially in the founding of *The British Journal for the Philosophy of Science*.

SYNTHESE

*An International Journal for Epistemology, Methodology and
Philosophy of Science*

Editor: JAAKKO HINTIKKA

University of Helsinki and Stanford University

Board of Consulting Editors: David M. Armstrong, Yehoshua Bar-Hillel, Arthur W. Burks, Robert S. Cohen, Donald Davidson, Dagfinn Føllesdal, Hans Freudenthal, Béla Juhos, Jerrold J. Katz, Imre Lakatos, G. Nuchelmans, Wesley C. Salmon, Wolfgang Stegmüller, Erik Stenius, Patrick Suppes, Klemens Szanlawski, H. Törnebohm, Marx Wartofsky.

Contents of Volume 18, No. 1, January 1968

The Structure of Logical Probabilities by *Jens Erik Fenstad*

The Whole Truth by *Richard C. Jeffrey*

On the Confirmation of Hypotheses about Regions of Existence by *Håkan Törnebohm*

Causal Connection by *William Ruddick*

A Generalization of the Principle of Causality, Which Makes It Applicable to Evolutionary

Systems by *Claude Hillinger*

Causality Requirements and the Theory of Relativity by *Peter Haves*

No more Discovery in Physics? Review Essay by *Joseph Agassi*

Reviews

Books Received

Annual Subscription Price Dfl. 60.—(\$17.00)

Personal Subscription Price Dfl. 35.—(\$9.80)

D. REIDEL PUBLISHING COMPANY
DORDRECHT-HOLLAND

VOLUME X

JUNE 1968

No. 1

RATIO

edited in co-operation with

H. B. Acton (London), P. Bernays (Zürich), A. Church (Princeton),
H. Cherniss (Princeton), G. Henry-Hermann (Bremen), W. C.
Kneale (Oxford), E. Nagel (Columbia), K. R. Popper (London)
and J. W. N. Watkins (London)

by

Stephan Körner (Bristol)

W. STEGMÜLLER

Towards a Rational Reconstruction of Kant's
Metaphysics of Experience, Part II

PETER WINCH

Wittgenstein's Treatment of the Will

JÜRGEN MITTELSTRASS

Predication and Recurrence of the Same

WOLFGANG YOURGRAU AND
CHARLES WORKS

A New, Formalized Version of the Verifiability
Principle

RICHARD S. WESTFALL

Huygens's Rings and Newton's Rings: Periodicity
and Seventeenth Century Optics

REVIEW:

GRAHAM BOND

Lewis White Beck: Studies in the Philosophy of
Kant

Price 15s. net

Annual Subscription: 27s. 6d. net

BASIL BLACKWELL · OXFORD

T H E O R I A

A Swedish Journal of Philosophy

VOLUME XXXIII

1967

PART 3

CONTENTS

Gunnar Aspelin: Historische Kausalsätze

Charles B. Daniels: Immediate knowledge: Ayer, Strawson and Shoemaker

Sven Danielsson: Modal logic based on probability theory

Jaegwon Kim: Psychophysical laws and theories of mind

William E. Mann: Definite descriptions and the Ontological Argument

Douglas Odegard: Arthur Pap and the Paradox of Analysis

Dag Prawitz: Completeness and Hauptsatz for second order logic

Irving Thalberg: Verbs, deeds and what happens to us

Gunnar Aspelin: 'Idea' and 'perception' in Locke's Essay

Charles F. Kielkopf: 'Ought' does not imply 'can'

CWK Glerup, Öresundsvägen 1, Lund, Sweden

Revue Internationale de Philosophie

Editorial Board: Profs. Ph. DEVAUX, H. GOUHIER, Ch. PERELMAN
and Mrs. Simone SERVAIS

Secretary: 99 avenue de l'Université, Bruxelles 5

Each number is devoted to a particular movement, a particular philosopher, or a particular problem. As complete a bibliography as possible is given.

Truly international in spirit, the *Review* chooses its collaborators among scholars of all countries. The articles are written in English, French, German, Italian, or Spanish.

Publishes 4 numbers annually. (Europe: 350 Belgian francs; other countries: 400 Belgian francs.)

Coming numbers for 1968

Philosophie du Langage au XVIII^e Siècle, Eug. Dupréel, L'Humanisme
L'Analogie

FOUNDATIONS OF LANGUAGE

International Journal of Language and Philosophy

Editorial Board: Morris Halle (MIT), Peter Hartmann (Münster/W), K. Kunjnni Raja (Madras), Benson Mates (Univ. of California), J. F. Staal (Amsterdam), Pieter A. Verburg (Groningen), John W. M. Verhaar (Nijmegen)

VOLUME IV, No. 1, JANUARY 1968

Articles: Wolfgang Motsch: Wolfgang Steinitz zum Gedenken — George Lakoff: Instrumental Adverbs and the Concept of Deep Structures — Paul Kiparsky: Tense and Mood in Indo-European Syntax — Sian L. Yen: Two Measures of Economy in Phonological Description — *Reviews* by: J. E. B. Gray, G. J. M. Bartelink, C. A. van Peursen, F. R. Palmer, Robert A. Hall, Jr., F. van der Blij, Henry M. Hoenigswald and Martin Hollis — *Publications Received*

FOUNDATIONS OF LANGUAGE is published quarterly

Annual subscription price /42.—(\$11.75)

D. REIDEL PUBLISHING COMPANY/DORDRECHT - HOLLAND

Philosophy and Phenomenological Research

AN INTERNATIONAL QUARTERLY

Edited by MARVIN FARBER

MARCH, 1968, VOLUME XXVIII, No. 2

CONTENTS

- I. ARTICLES
Moore and Ducasse on the Sense Data Issue, by PETER HARE and RICHARD KOEHL
Consciousness and its Correlatives: Eliot and Husserl, by JITENDRA KUMAR
A Phenomenological Model of the Artistic and Critical Attitudes, by JOSEPH BENSMAN and ROBERT LILIENFELD
Marxism and Ethics: A Methodological Inquiry, by A. JAMES GREGOR
The Reformulation of the Question as to the Existence of God, by HAROLD A. DURFEE
Requirements for the Validity of Induction: An Examination of Charles Peirce's Theory, by CHUNG-YING CHENG
- II. DISCUSSIONS
- III. REVIEWS, COMMUNICATIONS, NOTES AND NEWS, AND RECENT PUBLICATIONS

Subscription prices \$6.00 per year for individuals, \$7.50 per year for institutions
Separate issues, \$1.65 and \$2.00 respectively
Postage and handling for back issues only

A circular listing the contents of the Journal since 1940 will be sent upon request

Address all communications to

PHILOSOPHY AND PHENOMENOLOGICAL RESEARCH
State University of New York at Buffalo, Buffalo, New York, 14214

INQUIRY

An Interdisciplinary Journal of
Philosophy and the Social Sciences
Edited by ARNE NÆSS

Volume 11

Number 1

Spring 1968

Towards a Phenomenology of Morals

Kenneth I. Mills

Social Structures and Social Functions: the Emancipation of
Structural Analysis in Sociology

Filippo Barbano

Experience, Reality, and Conditions for Description

Lars Brink

Perception of Size

Age Slomann

Discussions:

Words and Tools Again

Leon Galis

Taylor on the Reduction of Teleological Laws

Joseph Margolis

Reply to Margolis

Charles Taylor

*Published quarterly by Universitetsforlaget, P.O. Box 307,
Blindern, Oslo 3, Norway*

Annual Subscription: 35s., \$6.00, postage included

"SCIENTIA"

INTERNATIONAL REVIEW OF SCIENTIFIC SYNTHESIS

(1968 - 62nd Year)

Scientific Committee

G. ABETTI - G. B. BONINO - P. CALDIROLA - L. CALIFANO - R. CANESTRARI -
G. COLONNETTI - B. DEFINETTI - G. MERLA - G. MONTALENTI - E. PADOA - G. SANSONE -
A. SESTINI - V. SOMENZI - S. TONZIG.

Editor N. BONETTI

"SCIENTIA" publishes original articles by the most famous scientists from
all countries, in all the fields of science and philosophy.
It is a magazine of synthesis, that is of convergency of the various branches
of knowledge.

It is a magazine of multi-specialistic information which promotes the
"bridge" disciplines, the most active in the progress of research.

It is for the formation of a culture the most adherent to the modern scientific
and philosophic requirements.

Published every two months, each issue containing approximately 200 pages.

Articles are published in the original language of the author i.e. English,
Italian, French, Spanish and German. Every issue contains a supplement
with a French translation of the non-French articles. Each issue contains:
Book reviews - Analysis - Chronicle.

SUBSCRIPTION RATES: Italy Lire 15,000 - Abroad \$32,000 or £13

Current issue: Italy Lire 3,000 - Abroad \$6,50 or £2.

Back issues: Italy Lire 500 - Abroad \$1,50 or 11s. (in stamps).

"Scientia" representative in U.K.: Atlas Publishing & Distributing Co. Ltd. -
334 Brixton Road - London S.W.9

"SCIENTIA" - Via Donizetti, 55 - Milano (Italy)

THE BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

Committee 1968-69

President: Mr J. L. Mackie

Vice-President: Dr G. J. Whitrow

Hon. Secretary: Mr R. Harré

Hon. Treasurer: Dr W. Mays

Hon. Secretary of the Northern Group: Dr N. Wetherick

Ordinary Members: Dr J. Giedymin, Mr A. E. Musgrave, Mr H. D. Sluga

EDITORIAL BOARD OF THE JOURNAL

Professor J. D. Bernal

Professor R. B. Braithwaite

Professor H. Dingle

Professor Sir John Eccles

Professor Sir Harold Jeffreys

Professor W. C. Kneale

Professor S. Körner

Professor M. Polanyi

Professor Sir Karl Popper

Earl Russell

Dr G. J. Whitrow

Mr L. L. Whyte

Professor J. H. Woodger

The British Journal for the Philosophy of Science is the official organ of the Society, whose purpose is described by the following statement from its Constitution and Rules :

‘The purpose of the Society is to study the logic, the method and the philosophy of science as well as those of the various special sciences, including the social sciences.’

The Society holds meetings, at which papers are read and discussed, once a month during the academic year. These are held, by kind permission, in University College, London.

Ordinary membership is open to all persons who have been proposed by two members, approved by the Committee, and elected at an ordinary meeting. The annual subscription is 20s. Those who wish to join should write to the Hon. Secretary at Linacre College, Oxford.

The Northern Group of the Society meets regularly at various centres in the north of England. Those who wish to join it should write to the Hon. Secretary of the Northern Group at the Undergraduate School of Studies in the Social Sciences, University of Bradford, Bradford 7.

Brit. J. Phil. Sci. Vol. 19 No. 1 May 1968

The British Journal for the Philosophy of Science

ARTICLES

- 1-25 NICHOLAS MAXWELL Can there be Necessary Connections
between Successive Events?
27-42 WILLIAM KNEALE Scientific Revolution for Ever?
43-56 FRED WILSON Definition and Discovery (II)

DISCUSSIONS

- 57-63 WILLIAM H. BAUMER Confirmation Still Without Para-
doxes
63-67 JEFFREY BUB AND MICHAEL RADNER Miller's Paradox of
Information
67-69 JON DORLING Length Contraction and Clock Synchroniza-
tion
70-71 L. JONATHAN COHEN A Note on Consilience
73-83 REVIEWS
85-87 ABSTRACTS
88-92 RECENT PUBLICATIONS

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, N.W.1.

American Branch : 32 East 57th Street, New York, N.Y. 10022

Single issues 20s. net in U.K. (\$2.75 in U.S.A.)

Subscription price 60s. net in U.K. (\$9.50 in U.S.A.)

Volume 19 Number 2

August 1968

The British Journal for the Philosophy of Science



CAMBRIDGE UNIVERSITY PRESS

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

Editors: MARY B. HESSE, D. H. MELLOR

Review Editors: B. M. FLOOD, G. M. K. HUNT

Subscriptions

The *Journal* is published quarterly (May, August, November and February). The subscription price is 60s. net in U.K. (U.S.A. \$9.50) per volume post free; single issues 20s. net in U.K. (U.S.A. \$2.75) plus postage. The subscription is 40s. (U.S.A. \$7.00) for members of the British Society for the Philosophy of Science, who should order direct from the publisher. Orders may be sent to any bookseller or direct to Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1. Orders from U.S.A. and Canada should be sent to Cambridge University Press American Branch, 32 East 57th Street, New York, N.Y. 10022, U.S.A. Second class postage paid at New York, N.Y.

Notes to Contributors

Articles and correspondence should be sent to the Editors, and books for review to the Review Editors, at the Whipple Museum, Free School Lane, Cambridge. Articles should not normally exceed 5,000 words, and should be accompanied by abstracts of about 100 words which will then appear in *Dialectica*, *Synthese* and *Philosophy of Science*.

All contributions should be typed in duplicate on one side only with wide margins and double-spaced. Footnotes, which should be as few as possible, should be typed separately, double-spaced at the end of the text. They should be referred to by consecutive numbering throughout the text, in which they will be inserted in page proof. Bibliographical references should be listed alphabetically at the end, and reference made to this list in the text rather than in footnotes. The style of references should in future follow the Harvard system.

Long quotations are printed in small type without quotes, and should be so marked in the typescript. Diagrams, tables, and illustrations should be on separate sheets, with their desired position in the text clearly indicated. Symbolic formulae appearing as separate lines should be clearly distinguished in the text, and foreign letters (Greek, German, etc.) indicated in the margin. On matters of style and layout, authors should consult T. W. Chaundy et al.: *The Printing of Mathematics* (Oxford University Press, 1964).

Offprints of articles should be ordered from the Editor when corrected galley proofs are returned. 50 are supplied free; the price of extra copies is obtainable from the publisher.

Advertising

Particulars of advertising in the *Journal* may be obtained from Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1.; or in U.S.A. and Canada, Cambridge University Press American Branch, 32 East 57th Street, New York, N.Y. 10022 U.S.A.

Explanation and Value Neutrality

by JAMES LEACH



I

Following the progressive elimination of teleology from the province of the scientist in the seventeenth century, the domains of science and values seemed finally isolated and immune each from the other. Value free scientific objectivity emerged as the ideal, with the very notion of scientifically supported values exploded as a 'naturalistic fallacy'. Finally, the convergence in value theory of intuitionists, emotivists and oxonians, conjoined with the success and influence of a positivist philosophy of science, consummated the separation of science and values. The issue seemed securely resolved. Particularly enthroned was the value neutrality thesis of science (VN) which prohibits the scientist from making value judgments.

Professor Ernest Nagel, in his illuminating treatise on the logic of explanation,¹ follows the lead of Carnap, Jeffrey and Hempel in opting for this thesis. Most attempts to disunify the empirical sciences, on the grounds that the natural sciences are value free while history or the social sciences are not, he disposes of nicely with an impressive barrage of counter evidence. In particular he correctly rejects the idealist claim that to explain an action one must appraise that action. Hence he rests the unity of the sciences on the exclusion of values.

Nevertheless, there remains one argument against which his barrage clearly misses the mark. And the survival of this argument suffices to revive the issue of value neutrality. For the smoldering cinders of pragmatism, kindled by recent successes in statistical analysis, flare forth to suggest an alternative defence of the unity of science. But now a unity grounded on the inclusion, not the exclusion, of value judgments or decisions.

A survey of recent philosophic literature reveals progressively increased agreement, I think, that the amount of evidence necessary to warrant the acceptance of any particular explanatory hypothesis is not fixed but instead depends on our varied purposes. However, it seems less generally recognised that sufficient evidence also depends on pragmatic costs concerning the significance of making a mistake when acting on the hypothesis in question.

Received 12.xii.67

¹ E. Nagel, *The Structure of Science* (New York: Harcourt & Brace, 1961), pp. 485-502.

Generally, I shall argue in what follows, the reasonableness or acceptability of probabilistic hypotheses necessitates appraisal according to pragmatic criteria, and hence surrendering the value neutrality thesis even in its weakest form. When such hypotheses occur in fields like history where the data is often meager and where plausible competing hypotheses abound, the requirements for acceptable evidence must be loosened considerably. In other words, the degree of rigidity of such requirements, as a function of these cost factors, helps to distinguish the various empirical disciplines. Hence the goal of such varied disciplines as history, law, and science is not merely truth for its own sake but truth modified by other goals and criteria. Before turning to an elaboration and defence of the argument surviving Nagel's criticism, let me first clarify the VN thesis itself.

II

Though favoured by a long history, the VN thesis remains ambiguous even as formulated in the most recent literature. Generally, the thesis demands the scientist's value scheme to be logically divorced from scientific standards of explanatory validity and reasonable belief. It requires scientific inquiry to be objective both in the sense that the scientist remain evaluatively neutral when validating or appraising the rational correctness of his explanations, and in the sense of yielding the same conclusion for all competent inquirers. In short, the scientist qua scientist must make no value judgments in what has been called the context of justification.

But so formulated, the thesis leaves unspecified what sorts of value the scientist should avoid and hence what type of neutrality to maintain. Since the values one advocates, seeks or avoids are correlated with goals or objectives, the question quickly reduces to a concern with the goals of scientific inquiry. Three different kinds of goal seem relevant. The scientist might seek, first, truth and nothing but the truth; secondly, truth qualified by such other independent theoretical or epistemic values as explanatory power, simplicity, predictive content and relief from doubt; or finally these in addition to other pragmatic values associated with acting on his beliefs.

The VN thesis can, accordingly, take three progressively weaker forms. In its strongest and least defensible version it prohibits the scientist from making any value judgments or decisions at all. The scientist must seek no goals, not even truth. Since he neither accepts nor rejects hypotheses, he has no need, qua scientist, to replace doubt by true belief. At most he assigns degrees of confirmation or probability to hypotheses so that others, decision makers, can decide as to the truth or falsity of the beliefs. Let us

call this strong version the 'odds maker' view of science and of the VN thesis. Should this position prove too strong, a weaker version of VN might then be advocated to the effect that the scientist, though himself a decision maker, seeks only one goal, has only one overriding value: truth at any cost. He is prohibited only from making any other type of value judgment than appraising hypotheses as to their likelihood to be true. Let us call this the 'truth seeker' view of VN. A still weaker, more liberal version is also possible. Many people construe science as a quest not only for true hypotheses but also for ones with additional cognitive, theoretical or epistemic values. Thus the scientist, on this 'cognitive decision maker' version of VN, can legitimately gamble with truth so long as the costs involved are purely theoretical. He is proscribed only from making value judgments of a practical sort based on the seriousness of consequences of acting upon his accepted hypotheses.

This liberal interpretation, of course, sets a limit to the VN thesis. For were any further practical values or goals to be countenanced within the methodology of science, little or nothing would remain toward which the scientist must be neutral. The force of the methodological imperative of VN would be negligible. However, one could interpret VN not as related to scientific goals but to subjective biases of the investigator. In this sense the scientist would be prohibited from taking his own values, preferences and attitudes into account when appraising his acceptances. But this seems to me a quite independent issue. For whether the scientist qua scientist has practical goals and hence whether or not he must be neutral toward some values is not in any way a function of whether these goals can be formalised, codified and made objectively uniform for all scientific inquirers. Whether one's personal values can be objectively defended is, indeed, a most important but open question. Nonetheless, it is independent of what kinds of goals scientists ought to pursue, and hence of the VN thesis, in any of the three versions marked off above, which is our concern here.

Thus to challenge VN in its weakest and most cogent form one must defend what we will call the 'practical decision maker' view of science by arguing that the scientist in the logic of his inquiry also makes pragmatic value judgments in pursuit of practical goals. This amounts to arguing that two inquirers, with the same evidence and making the same probability assignments, might nonetheless rationally disagree about the acceptability of a given explanatory hypothesis on the basis of attaching different values to the seriousness in action of making a mistake. The case against VN seems best argued, then, on the grounds of recent work in statistical analysis concerned with the problem of rational decision making in the face of uncertainty. This type of argument, advocated by such experimentalist

philosophers as Braithwaite, Churchman and Rudner,¹ builds on the earlier statistical work of Neyman, Savage and Wald in application to problems of quality control and more generally to problems of non-deductive inference.

Applications of statistical procedures are to be found, of course, in most of the natural and social sciences where a major goal is to decide what to believe. But they are also used in determining insurance rates and in operations research where the goal is to determine how to act. Hence, the basic problem of statistics, rational decision making in the face of uncertainty, suggests not so much a structural similarity but a *partial reduction* or dependency relation between these different kinds of problems: how to act and what to believe in the face of uncertainty. Accordingly, Professor R. Rudner has formulated what I take to be the most explicit and compelling attack against the value neutrality thesis by posing his argument in this context.

But since Rudner's case is incomplete as it stands and has subsequently been subjected to sustained counterattack in the recent literature by Professors Carl Hempel and Isaac Levi, it will be instructive to unpack the argument in some detail. We will then be in advantageous position to assess the Hempel-Levi defence of a value free science by locating the form of the thesis and the accompanying issues more clearly. In Rudner's argument four statements are explicitly noted as premises from which the conclusion is claimed to follow. Let us list them as follows:

- (1) 'The scientist as scientist accepts or rejects hypotheses.'
- (2) 'No scientific hypothesis is ever completely verified.'
- (3) Therefore, 'The scientist must make the decision that the evidence is *sufficiently* strong or that the probability is *sufficiently* high to warrant the acceptance of the hypothesis.'
- (4) The 'decision regarding the evidence and respecting how strong is "strong enough" is going to be a function of the *importance*, in the typically ethical sense, of making a mistake in accepting or rejecting the hypothesis.'
- (5) Therefore, 'the scientist as scientist does make value judgments.'²

From the above description of the three interpretations of VN we can see that each version denies the conclusion (5) on the basis of rejecting a different premise. The 'odds maker' view denies (1), the 'truth seeker' view rejects (3), and the 'cognitive decision maker' view denies (4). Hence if all of the premises can be sufficiently supported, the VN thesis will have to be abandoned.

¹ R. Braithwaite, *Scientific Explanation* (New York: Harpers, 1953); C. W. Churchman, 'Statistics, Pragmatics, Induction', *Philosophy of Science*, 15 (1948); R. Rudner, 'The Scientist Qua Scientist Makes Value Judgments', *Philosophy of Science*, 20 (1953).

² Rudner, p. 2.

III

Now to some elaboration. Premise (2) states a central tenet of empiricism to which all covering law theorists subscribe. In accord with the requirement of empirical testability, it formulates one version of the fallibilist claim that all scientific explanatory hypotheses are empirical and hence corrigible. No such statement is ever without risk or ever completely confirmed by any amount of evidence. With (2) thus undisputed by covering law theorists, the acceptability of (3) depends only upon that of (1). How then to defend (1)?

The issue here turns on how to interpret the function of the scientist. For to defend the value neutrality thesis by rejecting (1), i.e. to deny (5) by denying that the scientist accepts or rejects hypotheses, commits one to the 'odds maker' or 'guidance-counselor' view of science. On this rendering, espoused by Carnap¹ and earlier by Hempel², the scientist simply assigns degrees of confirmation to hypotheses relative to the total available evidence. He thus serves only as an advisor to policy makers who might want to apply such information to practical affairs. Carnap of course defines logical or inductive inference in this way, so that all probabilistic explanatory inferences in science consist not in an attempt to replace doubt by true or reasonable belief, but only in analytically assigning degrees of confirmation to hypotheses relative to the given evidence. Hence, if this strong construal of the VN thesis is convincing, the objection to (1), and thus to (5), concedes that the scientist qua policy maker must make the decision required by (3), yet insists that the scientist qua scientist, i.e. qua odds maker, merely determines the degree of probability for an hypothesis. The VN thesis would therefore dictate giving up (1). Determining how costly this price is and whether wisdom demands such a price then become our immediate problems.

No doubt the issue as to the aim and function of science is embarrassingly complex and difficult. At the present time we have no completely adequate answer. But Hempel's more recent analysis of explanation itself entails that the scientist do more than assign degrees of confirmation. Suppose we ask whether the covering law theory is intended as an explication of 'an explanation' or of a 'rationally acceptable explanation'? Clearly covering law theorists intend the latter. They have never been satisfied to analyse the notion of explanation merely in terms of a valid argument containing empirical laws among its explanans. Instead, they correctly insist that the explanans also be rationally acceptable in order to provide

¹ R. Carnap, *Logical Foundations of Probability* (Univ. Chicago Press, 1950).

² C. Hempel, 'Review of Churchman's Theory of Experimental Inference', *Journal of Philosophy*, 46 (1949).

good grounds for the explanandum, and hence to be a 'correct explanation'.

Further, when attempting to eliminate the inconsistency of inductive inference, Hempel clearly recognises the inadequacy of his earlier requirements for deductive systematisations as sufficient conditions of rationally acceptable explanations generally. He thus appeals for instance to the principle of total evidence. This immediately transcends the scope of inductive logic proper to the application or methodology of induction and to general pragmatic (as opposed to syntactic and semantical) considerations. Now, if the task were simply to explicate the notion of a formally sound (deductively or inductively) potential explanation, then pragmatic and methodological matters would be irrelevant: the scientist might escape as an odds maker by simply assigning degrees of confirmation to explanatory hypotheses. But since the covering law theory is designed to explicate the notion of an acceptable or reasonable explanation, which often requires acceptable statistical hypotheses as premises, the principle of total evidence and other methodological principles must be invoked. And this means that the total evidence, at least, must already have been accepted. In this case we must abandon the strong version of the VN thesis whereby the scientist pursues no values or utilities at all, not even truth. The scientist is at least a truth seeker and must decide when an hypothesis is to be accepted as true or rejected as false.

This suffices for the claim in premise (1) that an adequate explication or rational reconstruction of scientific procedures must include a statement to the effect that the scientist as scientist accepts or rejects hypotheses. Hence, if these considerations are cogent, (3) follows as a consequence. That is, the scientist must be seen not just as a counsellor of policy makers, as an odds maker, but as himself a decision maker. One of his major objectives is to replace doubt by true or reasonable belief, and to do this by deciding in the face of uncertainty when the evidence for an explanatory hypothesis suffices to warrant his belief in the hypothesis.

Moreover, if corrigibility or the chance of making an error were the only relevant factor for the decision required by (3), the scientist would never reach any decision. He would merely keep increasing the amount of evidence indefinitely. But if (1) is true, if the scientist must decide in the face of uncertainty, then, contrary to the 'truth seeker' view, some additional criterion or rules of acceptance needs be invoked in order to make a reasonable decision. Even in the pursuit of truth one must assess or attach utilities to the degree true answers are desirable and mistakes undesirable. To weight all mistakes equally comes to refusing to accept an hypothesis as warranted until all the evidence is both in and affirmative. Thus the second

version of VN, that the scientist makes utility or value judgments regarding truth only and pursues truth at any cost, must also be rejected. Further, since (5) clearly follows from (4) and with (3) now established, the tenability of Rudner's argument rests upon the move from (3) to (4), from the fact that scientists must make decisions to the fact that these decisions are a function of the seriousness of error relative to pragmatic goals.

IV

Unfortunately, Rudner offers no argument in the article cited for this cardinal connection. Instead, he fills the gap by appealing to examples from quality control and to some guiding theories of statistical inference. 'If', to use one of his cases:

the hypothesis under consideration were to the effect that a toxic ingredient of a drug was not present in lethal quantity, we would require a relatively high degree of confirmation or confidence before accepting the hypothesis—for the consequences of making a mistake here are exceedingly grave by our moral standards. On the other hand, if say, our hypothesis states that, on the basis of a sample, a certain lot of machine stamped belt buckles was not defective, the degree of confidence we should require would be relatively not so high.¹

Taking such an example as a paradigm case of scientific inquiry, he follows the statisticians by assimilating cases of deciding what to believe to cases of deciding how to act in the face of uncertainty. He concludes that the seriousness of making a mistake, reflected in the cost or value factor, must be weighed in any scientific assessment of statistical explanatory hypotheses.

What is unfortunate in the appeal to examples at this pivotal juncture in the argument, however, is not the lack of better grounds for moving from (3) to (4), for there are indeed better grounds. Rather, the suppression of the needed assumptions has led some defenders of the VN thesis to serious misunderstandings. As a result they reconstruct Rudner's argument by appealing to an altogether unnecessary premise to fill the gap, and then proceed to challenge this very premise. Professor Isaac Levi, for instance, clearly sees the required logical connection between believing or accepting an hypothesis and acting on the basis of the hypothesis relative to some specific objective or goal. This stresses again the need to interpret explanatory hypotheses, at least in part, methodologically as means or instruments adopted to achieve some specified objectives. Rudner's argument assumes, with Hempel and Levi concurring, that all scientific problem solving is a species of purposive behaviour. The assessment of the rational acceptability of explanatory hypotheses is consequently relativised to goals which

¹ Rudner, p. 2.

require weighing cost or utility factors in addition to degrees of probability or confirmation. The relevance of decision theory to the logic of inductive inference therefore rests in part at least on its applicability to inductive behaviour.

But when Levi states explicitly the assumptions necessary to move from premise (3) to (4) and hence to the denial of the VN thesis, (5), he switches issues. He commits Rudner and Churchman to two additional premises, one of which is unnecessarily strong. He nevertheless makes it the centre of controversy. The two additional assumptions cited by Levi are as follows:

- (6) 'To choose to accept a hypothesis H as true (or to believe that H is true) is equivalent to choosing to act on the basis of H relative to some specific objective P.'
- (7) 'The degree of confirmation a hypothesis H must have before one is warranted in choosing to act on the basis of H relative to an objective P is a function of the seriousness of the error relative to P resulting from basing the action on the wrong hypothesis.'¹

Now (7) he accepts on the authority of the statistical theories of Pearson, Neyman and Wald. But since (7) without (6) obviously fails to yield (4), the tenability of (4) is made to rest on (6), on some sort of identity or equivalence between believing an hypothesis and acting on the basis of it relative to some specific objective P. For, as Levi queries, 'Otherwise how can the role of value judgments in determining minimum probabilities be explained?'

But surely a weaker commitment suffices in response to this question. Whether or not Rudner and Churchman would accept (6) as an adequate formulation of their suppressed premise seems simply irrelevant to the issue. Hence Levi and others raise a false issue and thereby confuse what I think are the important ones. Let me clarify. Levi assumes that the desired conclusion presupposes a 'behaviouralist' construal of belief whereby 'accepting a proposition H as true' is synonymous with, equivalent to, equated with or reducible to 'acting on the basis of H relative to a practical objective P'.

However such an assumption is not essential to yield (4). A behaviouralist view of belief, in other words, is unnecessarily strong. The only conditions required for the argument are that belief be logically tied to action in such a way that beliefs can be *partially justified* by practical considerations of action, and that different mistakes in action be taken with

¹ I. Levi, 'Must the Scientist Make Value Judgments', *Journal of Philosophy*, 67 (1961), 348. For additional instances of invoking (6) in a similar context see the discussion by Kyburg, Black, Sellars, Jeffrey, Levi and Bar-Hillel of a paper by Braithwaite in Kyburg and Nagel (Eds.), *Induction: Some Current Issues* (Conn: Wesleyan Univ. Press, 1963), pp. 196-204.

² Levi, 'On the Seriousness of Mistakes', *Philosophy of Science*, 29 (1962), 50.

different degrees of seriousness. And these conditions, satisfied by a critical cognitivist view as well as by a behavioural account of belief, assure that the acceptance of an hypothesis is not merely the result of a quest for a single value or goal, truth and nothing but the truth. The logical connection between belief and action would, on the critical cognitivist view, be considerably weaker than identity, deductive entailment or even material equivalence. It might in fact be as weak as nomological implication even in its defeasible version, i.e. with a *ceteris paribus* clause. Hence critical cognitivism contrasts with behaviouralism by viewing action or behaviour merely as evidential manifestations, signs, criteria or contingently necessary conditions of belief. Thus in place of the equivalence condition in (6) two nomological implications suffice:

- (6') To choose to accept a hypothesis H as true implies as a contingently necessary condition the disposition to act on the basis of H relative to some specific objective P.
- (6'') To be disposed to act on the basis of H relative to some specific objective P implies as a contingently necessary condition acting on the basis of H relative to P in the appropriate circumstances C.

Once these replace (6), the issue no longer turns on a behaviouralist reduction of belief to action but clearly focuses on the questions posed by Hempel's attempt to reinforce the principle of total evidence by formulating acceptance rules of application for rational belief. The issue concerns as we have seen, whether the scientist pursues one or more goals and what kind of values the assigned utilities are to represent, e.g. pragmatic or just cognitive utilities.

Once again, however, confusion threatens. For Rudner's premise (4), invoking the notion of the seriousness or importance of error, does not commit one to a purely instrumentalist view. It does not make science merely a handmaiden of practice. Nor does it require pragmatic costs or utilities to replace truth, high confirmation, simplicity, or even explanatory power as appropriate criteria for appraising the reasonableness of hypotheses. Nor does it entail defining any of these criteria in terms of pragmatic utility. For these can be interpreted as independent weights, each important for an adequate theory of inductive inference and for the assessment of rational choice among explanatory hypotheses. Yet no criterion is either a necessary or sufficient condition of the reasonableness of a belief or, more precisely, of a decision to accept a belief. The proper relative weights to be assigned these criteria, moreover, remains a difficult and open question.

The burden of Rudner's argument, then, is to show that pragmatic costs enter into the scientist's decision to accept or reject hypotheses. Thus the move from (3) to (4) requires the suppressed premises (6'), (6'') and (7) to

provide the defeasible link between the decision to accept H and actions performed upon H in pursuit of pragmatic goals. This amounts to resting the reasonableness or rational acceptability of H relative to P not only on whether H is true or has high epistemic utility, but also upon its pragmatic value, upon whether or not action based upon H is warranted.

One further caution. The point is not that if the appropriate act is unwarranted, the disposition and belief leading to the act are thereby unreasonable. For the epistemic interdependence of acts, dispositions and beliefs makes the appropriate logical connection between these notions much more complex than this stronger claim allows. Rather we argue the weaker claim: the unwarrantedness of the action is but one determining factor in appraising the disposition and belief. No doubt the conceptual meanings of 'belief' and 'action' and hence the adequacy of (6') and (6'') still await clearer analysis. Almost inconceivable to me, however, is that an adequate analysis of 'belief' could be provided which is logically *unrelated* to 'action'. For there would then seem to be no way to account for the important influence of beliefs on action. Hence, if 'X believes H, relative to P', in conjunction with other statements, defeasibly implies 'X performs act A, relative to P' then it seems reasonable that the appraisal of A should bear on the appraisal of the decision to believe H. This I take to be the crux of Rudner's case.

v

Having warded off the false issues concerning the move from (3) to (4), a behaviouralist reduction of belief to action and an instrumentalist reduction of truth to utility, let us return to the main issue at hand. Suppose, then, an inquirer to be seeking the truth and nothing but the truth, to be concerned only with adding true statements into the body of scientific knowledge K. To be rational, in the second version of VN, the inquirer must weight each possible mistake he might make as equally serious. But such a condition cannot be satisfied on Rudner's argument since different mistakes have different degrees of seriousness or importance for scientists pursuing varied goals. This conflict serves as a basis to continue the defence of (4). It will also be helpful at this juncture to see how Hempel develops his notion of 'epistemic utility' in defence of the 'cognitive decision maker' view of science.

The problem of establishing acceptance rules for scientific explanatory hypotheses, of deciding which scientific hypotheses to accept and thus add to the body of scientific knowledge, turns on what kind of values the utilities assigned to outcomes are to represent. Hempel maintains that:

the utilities should reflect the value or disvalue which the different outcomes have from the point of view of pure scientific research rather than the practical advantages or disadvantages that might result from the application of an accepted hypothesis, according as the latter is true or false. Let me refer to the kind of utilities thus vaguely characterized as *purely scientific, or epistemic, utilities*.¹

Beyond the need to clarify this distinction between pragmatic and epistemic values, the problem concerns finding a measure of the epistemic utility of adding an hypothesis h to the previously established system of knowledge K . Carnap's rule of maximising estimated utility warrants the acceptance or rejection of h as epistemically rational in accordance with the following rule:

'Accept or reject h , given K , according as $c(h, K) > \frac{1}{2}$; or $c(h, K) < \frac{1}{2}$; when $c(h, K) = \frac{1}{2}$, h may be accepted, rejected, or left in suspense.'²

Since on this account a scientist is interested only in accepting h when it is true, the possible correct answers of accepting h when true and rejecting h when false are considered equally desirable and the corresponding mistakes equally undesirable. Hence, we are led to the unsettling recommendation of accepting, and hence of acting upon the basis of, h when its degree of confirmation or probability is merely 0.51.

The obvious difficulty with Carnap's account, beyond leading to paradox, is that the rule is much too liberal to be suitable for even pure scientific procedure. Further, Levi notes the same problem when other methods are used instead of Carnap's maximising utility (e.g. Bayes method and the method of significance testing proposed by Neyman and Pearson). He also acknowledges that accepting h 'when its degree of confirmation is low does not seem reasonable unless this acceptance is reduced to action undertaken to realise some objective other than seeking the truth and nothing but the truth'.³ Levi, however, pursues the third option, suspended judgment or remaining in doubt, as a way of avoiding the difficulty and hence of defending the view that scientists accept or reject hypotheses in quest of truth and nothing but the truth. In this sense, the scientist suspends judgment on h when he feels the total available evidence warrants neither the acceptance nor the rejection of h .

Consider one of Levi's examples of the Bayes method applied to the problem of replacing doubt by true belief. Suppose an experimental

¹ C. Hempel, 'Inductive Inconsistencies', *Logic and Language* (Dordrecht-Holland: Reidel Publishing Co., 1962), p. 154. Cf. also Hempel's most recent statement in 'Recent Problems in Induction', R. Colodny (Ed.) *Mind and Cosmos* (Pittsburgh: University of Pittsburgh Press, 1966).

² C. Hempel, 'Deductive Nomological vs. Statistical Explanation', *Minnesota Studies in the Philosophy of Science*, iii (Minnesota: University of Minnesota Press, 1962) 155.

³ I. Levi, 'On the Seriousness of Mistakes', p. 55.

psychologist wants to determine whether or not some person has extra-sensory perception. By assuming that the subject has ESP if he guesses correctly the colours of cards drawn randomly with a frequency greater than 0.6, and otherwise not, the psychologist decides on the basis of a sample of guesses whether the long range frequency of correct guesses is 0.6 or less (H_1), or is greater than 0.6 (H_2), when A_1 is the act of accepting hypothesis H_1 .

Since the experimenter presumably is interested only in accepting the hypothesis as true which is true, he should consider the possible correct answers O_{11} (accepting H_1 when it is true) and O_{22} (accepting H_2 when it is true) as equally desirable and the corresponding 'mistakes' or 'errors' O_{12} and O_{21} as equally undesirable. . . . Consequently, the matrix for this problem would be as follows:

	H_1	H_2
A_1	1	0
A_2	0	1

The Bayes method recommends adopting A_1 if the probability of H_1 is greater than $\frac{1}{2}$ and A_2 if the probability of H_1 is less than $\frac{1}{2}$. Consequently, if the probability of H_2 . . . is 0.51, we would be warranted in accepting H_2 (adopting A_2).¹

Now, to avoid this unsettling and unreasonable consequence, yet to preserve application of the method to cases where the goal is to seek the truth and nothing but the truth, Levi revises the problem by adding a third option: the act S of suspending judgment or remaining in doubt. In this case the matrix changes to the following:

	H_1	H_2
A_1	1	0
A_2	0	1
S	k	k

Here the utility k is the value of act S when the hypothesis H_1 is true. The strategy consists in assigning a sufficiently high value to k so that a high degree of confirmation is required to warrant the acceptance of any hypothesis. But the price of thus avoiding the unsettling consequence amounts to selecting arbitrarily the value assigned to k .

VI

However, this arbitrary assignment brings us back to the crux of Rudner's move from premise (3) to (4), from the fact that scientists must make decisions (which Levi and Hempel concede) to the fact that these decisions are a function of the seriousness of error. For Rudner's argument hinges on the assumption that any arbitrary selection of a value assigned to k in such

¹ I. Levi, 'On the Seriousness of Mistakes,' p. 54.

cases as the above would be unreasonable. Particularly so if there are grounds in action for making such assignments of values. This is why premises (6') and (6''), which connect beliefs or the acceptance of hypotheses with acting upon the basis of them relative to some goal, provide the crucial link between the decision to accept in (3) and the different degrees of seriousness or error in (4). Instead of assigning a value to k arbitrarily or on the basis of sheer convention, we take our cue from (6'), from the fact that belief implies a disposition to act. Our choice of a value for k then becomes a function of how serious a mistake would be if we acted upon the hypothesis, i.e. if we accepted a false hypothesis or rejected a true one. This in turn depends upon how useful the hypothesis in question is to us relative to our goals. So, while on purely logical grounds the value assigned to k is arbitrary, still a rational choice of assignments can and must be made on extra-logical grounds, viz. on the basis of purposiveness and pragmatic utilities. Relevant to this choice are the purposes for which we intend to use the hypothesis, and the relative advantages and disadvantages of acting upon it relative to these objectives.

It should be noted that Levi is not unaware of this type of consideration. After noting the unhappy result, the arbitrary assignment of a value for k , he offers further revision of his non-behavioural decision-theoretic approach to scientific inference by introducing the notion of 'degrees of caution'. In the case of his truth-seeking experimenter, we have a scientist who 'takes mistakes more seriously in relation to eliminating doubt' than might some others. Unlike a more tenacious scientist, he exercises a higher degree of caution, indicating 'that to seek truth and nothing but the truth is to aim at a complex objective. It is an attempt to eliminate doubt tempered by an interest in finding truth and avoiding error.'¹ But this of course opens the door to the move from premise (3) to (4). It means, as Levi himself acknowledges, 'that the choice of a problem whose goal is to replace doubt by true belief and, as a consequence, the choice of a degree of caution in realizing such a goal can be both influenced and justified by practical considerations'.²

Moreover, the fact that this conclusion implies nothing about any kind of behavioural reduction constitutes less solace for the value neutrality theses than Levi recognises. Nevertheless, in concentrating his effort against (6) Levi fails to see that by granting that the scientist must reckon with the seriousness of mistakes, though not a reductionism of belief to action, he has nonetheless conceded that different mistakes need to be taken with different degrees of seriousness. In other words, he totally neglects (6') and (6'') which together suffice to get Rudner's argument to the conclusion (5),

¹ I. Levi, 'On the Seriousness of Mistakes,' p. 56.

² *Ibid.* p. 57.

that scientists qua scientists make value judgments regarding the seriousness of errors in acting upon accepted or rejected hypotheses. For, the different purposes for which one acts clearly require different cost factors and, as a result, different degrees of seriousness. Hence Levi in effect concedes (5), the denial of VN even in its weakest form.

Clearly then, if the preceding considerations are cogent, the scientist must take into account in his decision to accept or reject hypotheses not only epistemic but also pragmatic utilities. Consequently the decision will vary with the kind of action to be based upon the hypothesis. Consider again the example from applied science cited earlier regarding the toxic ingredient of a drug. Because of the different utilities involved, a decision rule might warrant, on the same evidence, that the hypothesis in question should be accepted if applied to experimental animals, but rejected if applied to humans where an error would obviously be more serious. Hence the decision to accept or reject hypotheses is in part instrumental to action; it is partially a decision to adopt one of alternative courses of action. And this decision can be rationally defended only by considering the pragmatic gains and losses attached to the possible outcomes of the actions. This amounts to resting the acceptance of an hypothesis partially upon whether or not it leads to warranted action.

Having thus unpacked and defended the experimentalist argument opposing the varied versions of the value neutrality thesis of covering law theorists, let me now present it in full array:

- (1) The scientist as scientist accepts or rejects hypotheses.
- (2) No scientific hypothesis is ever completely verified, but all are corrigible.
- (3) Therefore, the scientist must make the decision that the evidence is sufficiently strong to warrant the acceptance of the hypothesis.
- (6') To choose to accept an hypothesis H as true (or to believe that H is true) implies as a contingently necessary condition the disposition to act on the basis of H relative to some specific objective P.
- (6'') To be disposed to act on the basis of H relative to some objective P implies as a contingently necessary condition acting on the basis of H in the appropriate circumstances C.
- (7) The degree of confirmation that a hypothesis H must have before one is warranted in choosing to act on the basis of H relative to an objective P is a function of the seriousness of the error relative to P resulting from basing the action on the wrong hypothesis.
- (4) Therefore, the decision in (3) regarding the evidence and respecting how strong is 'strong enough' to rationally accept H is going to be a function of the importance or seriousness, in the typically ethical sense, of making a mistake in accepting or rejecting the hypothesis.
- (5) Hence, the scientist as scientist does make pragmatic value judgments.

VII

Let me consider briefly one final but important objection to the foregoing argument, an objection posed most recently by Hempel, Nagel and Levi, but the roots of which lie deeply embedded in Aristotle's distinction between the theoretical and practical sciences. The experimentalist argument, it would seem, forces a reconsideration both of the meaning of scientific and historical objectivity, and of the relations between theoretical, technological and policy making aspects of rational inquiry. The issue of historical objectivity will be considered on another occasion.¹ Let me accordingly limit discussion to one related point.

Though Hempel is unable to find any 'satisfactory general way of resolving the issue between the two conceptions of science', the cognitive and pragmatic, he still claims 'that it would be pointless to formulate criteria of applicability by reference to pragmatic utilities; for we are concerned here with purely theoretical (in contrast to applied) explanatory and predictive statistical arguments'.² Nagel follows suit by objecting that the experimentalist argument misleadingly 'suggests that alternative decisions between statistical hypotheses must invariably lead to alternative actions having immediate practical consequences upon which different special values are placed'.³

This kind of objection forces one additional clarification. For the argument against the VN theses in no way depends on considering in the weighting of pragmatic utilities, only the 'alternative actions having *immediate* practical consequences'. No doubt some applied historical and scientific hypotheses will be closely linked logically to such actions. But others, of a higher degree of 'theoretical purity' will be more tenuously linked to such actions via the former ones. Despite the tenuousness of this linkage, it seems as ill-conceived to disregard the purposive and practical dimension of science as to disregard its empirical dimension. Surely there is an important sense in which the experimental data garnered from laboratory tests are logically related, as confirmation or disconfirmation, to 'purely theoretical' hypotheses. But then the practical considerations involved in experimental design of tests would bear heavily on the criteria of acceptance for even such theoretical hypotheses. Moreover, this linkage need be no more tenuous than the linkage of empirical significance between hypotheses containing theoretical constructs and relevant observational statements. And as it might be argued in Quinean fashion both that no observational statements are quite free of theory, and that no theoretical statements are

¹ J. Leach, 'Historical Objectivity and Value Neutrality', submitted to *Inquiry*.

² Hempel, 'Deductive-Nomological vs. Statistical Explanation', p. 162.

³ Nagel, *The Structure of Science*, p. 497.

quite free of experimental considerations, so it might also be argued in extension that no hypotheses are quite free of practical consequences.

In other words, it seems to make as much sense to speak of degrees of practical consequences and associated pragmatic utilities, with a system or network of hypotheses linked at its edges to immediate practical consequences, as it does for Hempel to speak of degrees of testability or empirical significance, with a network of hypotheses linked at its edges to immediate observations. Surely nothing in either linkage precludes the other. It should also come as no surprise that the goals or objectives used to weight the consequences of wrong estimations in the 'purely theoretical' sciences will be of a more complex nature and will require more general criteria than those used in the more practical sciences such as quality control and engineering. Hence, even if there are cases where the choice of policy for deciding to accept or reject hypotheses does not depend directly upon immediate practical consequences, still the costs or pragmatic utilities will be relevant in a more indirect manner. This should serve to dethrone the traditional VN thesis even in its weakest form.

College of William and Mary and University of Western Ontario.

Robots, Consciousness, and Programmed Behaviour¹

by KEITH GUNDERSON

I

Might we after all be a kind of robot? or might certain sorts of robots after all be a kind of us? Could a robot which simply did or had what it was supposedly programmed to do or have, be a robot which really felt and was conscious, had thoughts and intentions, deliberated and decided?

In 1748 in his book *L'homme machine* La Mettrie wrote:

We are in the position of a watch that should say (a writer of fables would make the watch a hero in a silly tale): 'I was never made by that fool of a workman, I who divide time, who mark so exactly the course of the sun, who repeat aloud the hours which I mark. No! that is impossible.'²

I shall apply the label 'defenders of La Mettrie's watch' to those who claim it could never be shown that our own mental life and behaviour is watch-like or programmed and that if a subject's mental life or behaviour is watch-like or programmed, then it cannot be like ours. From Descartes to Cybernetics one can tick off defenders of La Mettrie's watch. In his *Principles of Philosophy* Descartes wrote:

We do not praise automata for precisely carrying out all the movements for which they were designed, since they carry them out by necessity; we rather praise the maker for fashioning such precise machines, because he fashioned them not by necessity but freely.³

And about three hundred years later Paul Ziff in his provocative little article 'The Feelings of Robots'⁴ wrote:

MacKay has pointed out that any test for mental or any other attributes to be satisfied by the observable activity of a human being can be passed by automata.

Received 28.i.68

¹ Earlier versions of this paper were read and discussed with philosophers at the University of California, Santa Barbara, and at Clairmont College. A truncated version was presented at *The Chapel Hill Colloquium in Philosophy* (25 November, 1967). I am indebted to the participants in these discussions for helpful suggestions.

Some of the points developed here were suggested in my 'Interview with a Robot', *Analysis*, 23 (1963), 136-42.

² *L'homme machine*, by Julien Offray de La Mettrie, translated as *Man a Machine*, by Gertrude Carman Bussey (Open Court, 1953), p. 145.

³ In *Descartes-Philosophical Writings*, trans. by Elizabeth Anscombe and Peter Thomas Geach (Thomas Nelson and Sons Ltd., 1954), p. 188.

⁴ *Analysis*, 19 (1959), 67-68.

And so one is invited to say what would be wrong with a robot's performance. Nothing need be wrong with either the actor's or a robot's performance. What is wrong is that they are performances.

He then goes on to suggest reasons (similar to Descartes') why 'no robot could sensibly be said to feel anything':

Because there are not psychological truths about robots but only about the human makers of robots. Because the way a robot acts (in a specified context) depends primarily on how we programmed it to act. Because a robot could be programmed to act like a tired man when it lifted a feather and not when it lifted a ton.

So too Jonathan Cohen in his article 'Can there be Artificial Minds?'¹ marshalls a variety of arguments to support his contention that if a husband came to know and predict the future behaviour of his wife in the way in which we can come to control and predict (in principle, at least) the future 'behaviour' of a computer (with or without randomising elements):

. . . we should then say, I think, that his wife had no mind of her own, that her husband had already explicitly or implicitly done her thinking for her, and that the notion of programming an ideal mind, was therefore self-contradictory.

This conclusion will be of interest to us later. First, however, we shall be concerned with the jump Cohen makes from it to the following claims:

What I have been trying to point out is an opposition between the still familiar concept of mentality and the concept of total subservience to known or knowable rules. The metaphor of 'artificial minds' clearly destroys this opposition by ascribing mentality to programmed artifacts. That is why such a metaphor is—outside the laboratory—not merely eccentric and novel, but also undesirable. If we blur our concept of mind in this respect, then, however completely human beings were regimented in thought and action by their government's, party's, or church's ideological programme, it would be false to say that they behaved as if they had no minds of their own. They too could still be said to think for themselves. And that would be a shift in the notion of having a mind of one's own and thinking for oneself of which George Orwell's *Newspeak* could have been proud.

He goes on to write:

But I do not see how our cyberneticist's unfortunate wife could ever recover a mind of her own. Her husband's difficulty—suppose he wanted such a recovery—would be logical, not technical. Whatever randomising elements he introduced into her brain, or however much he asked other cyberneticians to programme her and not to tell him what they had done, neighbours who were 'in the know' would still nod their heads at one another as she walked past and murmur, 'Poor girl, she can never again think for herself or have any emotions that are really her own.'

Such arguments and attitudes have assumed a variety of forms and phrasings in recent years. But it will not matter if we neglect some of these

¹ In *Analysis*, 16 (1965).

nuances. For I intend to develop distinctions which will show that it would *never* follow from the fact of a subject being programmed that the subject failed to have thoughts, feelings, or intentions, perform this or that, etc. In so far as (some) robots can be shown to lack minds it will be for different reasons.

I am more interested in *why* the defence of La Mettrie's watch fails than in *that* it fails. *Why* it fails has a bearing on current issues in the philosophy of mind and philosophical psychology, namely the relationship between consciousness and behaviour and the explanatory limits of computer stimulation of cognitive processes.

I have elsewhere¹ argued against what I take to be the most influential² bad argument for saying that machines (or robots) could feel, think, or do such-and-such. I shall here argue against what I take to be the most influential bad argument for saying that machines (or robots) could *not* feel, think, and so on.

II

First consider a typically whimsical philosophical preamble to a discussion of the mentality of robots. Suppose a skull is cracked open and some micro-wires etc. fall out. We become suspicious. We inspect closely the subject we took to be the owner of that skull. Very interesting. Finally we exclaim: 'A robot . . . a programmed robot!' So we find *it* out. Whatever that subject had done in the past, whatever we had thought about it, 'his' mental life however we knew it is now unveiled as no more than the exhibited capacities of a programmed robot. Things are seen in a new light, and some would say it is now clear that at best the robot only 'behaved' and was 'conscious'.

So suppose we come to know that at one time so-and-so manufactured this robot, made this teeny mechanical brain which from the outset was endowed with a highly self-corrective micro-programme (self-corrective

¹ Keith Gunderson, 'The Imitation Game', *Mind*, 73, n.s. (1964), 234-45; reprinted in *Minds and Machines*, ed. Alan Ross Anderson, pp. 60-71.

² Note Denis Thompson's recent article 'Can a Machine be Conscious', where he refers to this argument as 'probably the most common argument in the machine debate', this *Journal*, 16 (1965), 36. And compare as well Michael Scriven in his 'The Compleat Robot: A Prolegomena to Androidology' where he begins Section 5 with a 'quote' from the air: 'Machines only do what we tell them to do, they are incapable of genuinely original thought.' Scriven's objections to this view mark an about face from the stance he took in 'The Mechanical Concept of Minds', in *Minds and Machines*, edited by Alan Ross Anderson, a fact often ignored by recent critics of that paper. What I say in this paper is, at least in conclusions, much in accord with Scriven's later views. Also Turing, Putnam, Minsky, Armer, *et al.* have at least sketched some dissatisfactions with the defence of La Mettrie's watch, and some of the things I say may be viewed as an elaboration of why such dissatisfaction has been justified.

in the way that certain programmes for digital computers are, so that the machine may alter its routines to avoid previous blunders in a chess game) and 'clothed and masked' it in such a way that it was 'virtually indistinguishable from men in practically all respects: in appearance, movement, in the utterances it uttered, and so forth'.¹ Now if we came to know such things, we might indeed with good reason think it necessary to revise some of our past attitudes towards and judgments about the subject who is now revealed as a robot. We would now suspect that certain descriptions and ascriptions which we had formerly thought appropriate of and to that subject were no longer appropriate. But exactly which ones, and why?

Give or take a few details the preamble flickers out and a set of philosophical questions are posed. The central problem usually turns out to be the one of trying to determine which psychological predicates would no longer be applicable to the robot and which ones would remain essentially 'intact'. Let me call this *the repredication problem* for the unmasked robot. There is obviously no obvious solution to the repredication problem for robots given only the foregoing description of the unmasking. For this is a highly flexible Halloween surprise which could be satisfied by many different kinds of robots. Rather like cars and rather unlike people, robots come in a variety of makes and models. What we could say truly about a conventional gear-change may not extend to an automatic.

Questions which pop up at this point are 'Is the programmed robot like a man under hypnosis? Or like a cunning dissembler? Or more like an actor with a script?' To assess these (often proffered) comparisons may be instructive. We find, for example, that the repredication which should occur in connection with the robot is very unlike the repredication which would be forced upon us were we to discover a subject to have done what he did as a result of being hypnotised (as, for example, where a man is hypnotised to insult someone who upon finding this out no longer feels insulted). Nor would it be like the case of the typical someone upon whom light breaks thus: 'Look here, now, you've been deceived all along. She's only been pretending. She doesn't really love you.' For to find out about the programme in the robot's 'life' (or life) is presumably, to find out something new about *all* the robot's 'life' (or life). But what would it be to find that someone had been hypnotised for the whole of his life, or had pretended for the whole of his life? What would it be for anyone to live an *entire* life under hypnosis, or an *entire* life of pretence, or have someone's *entire* life shown to be performance? Compare: '... eight, nine, ten', snap: 'You're no longer Descartes. You're just your old self.' But *what* or *which* old self? Nor could one pretend to be Descartes all of his life. For when

¹ Paul Ziff, this *Journal*, 16 (1965), 64.

would the pretence begin? In the womb? Pretence, hypnosis, performance, though different from each other, are alike in this, that they all demand settings. They demand an environment in which they can occur and with which therefore they can be contrasted—a backdrop of non-pretence, non-performance. They must *enter into* the subject's life at a certain point, at a certain time. They cannot constitute the subject's mode of life at all points, at all times. The case of the robot and his programme does not demand this to the same extent, which is one reason why Ziff is wrong (above) in his likening a robot acting according to a programme to an actor's performance. Unlike the actor with a script, the robot we've imagined can easily be seen as always having been a programmed robot, which insofar as its actions are concerned has performed them in accordance with that programme.

On the other hand it should be mentioned that only subjects with certain sorts of conscious capacities to begin with could be hypnotised, could be dissemblers, or performers with scripts. In other words consciousness is presupposed by abilities to act in accord with certain kinds of directives. It is not a byproduct of such directions. What this reveals in connection with robots is that it would be wrong to suggest that a robot could be made conscious as a result of a foxy programme. There is, to begin with, a subject, the robot, which has a programme. There is not, to end with, a subject, the programme, which is the robot. The importance of this morphology of mechanism is that what a programme can be made to make a robot do will in part depend upon facts about the robot's nature, its hardware, its input potential, which are independent of its programme. To assess *in toto* the mental machinations of a robot we need to consider not only how the robot is programmed, we need also to consider how and in what the programme is roboted.

If the robot so loosely (but typically) characterised at the outset was neither conscious nor unconscious but non-conscious then indeed we would have been deceived in any Buber-like 'I-thou' relationship we had thought existed between us and that robot. Here the appropriate repredication would be more analogous to that which would occur were we to discover a ventriloquist had been deceiving us with his dummy. For example: 'Charlie McCarthy seemed to be my most sympathetic friend until I found out he was made of wood. Fantastic. I felt and thought he had thoughts and feelings.'

But if the robot had certain basic capacities to begin with, then the case of discovering him (it) to be a robot might well not be a case where it would be appropriate to jump back and exclaim anything at all. If finding out about the robot is like finding out that a dog is not a dog but a wolf we may

well feel that it would not be polite to exclaim (rudely) in front of the robot. Just as if we desist, for moral reasons, from beating dogs, we may desist, for the same (good) reasons, from beating wolves.

III

But here let me anticipate disappointments with what I have so far said. The strategy underlying most discussions of minds and machines is designed to uncover whatever conditions must obtain before it would be appropriate to ascribe feelings, thoughts, and so on to robots. This is done with an eye to tampering with our all too unclear perspectives with respect to human minds and their underlying makeup. But to be told that we should not withdraw predications of feelings, thoughts, and so on to robots *if* the robots had certain conscious capacities to begin with is not, it must seem, to be told anything illuminating. Yet I think it is—particularly in the context of recent philosophical discussions of mind and machines and computer simulation of cognitive processes.

Once the distinction is made between a robot's basic capacities, its input potential, on the one hand, and its programme which may utilise these capacities on the other, the following claims seem in order:

If a subject is such that certain kinds of directives, commands, or controls (i.e. programmes), could not be imposed upon him (it), then, of course, the subject will not be able to perform various mental acts, infer this from that, prove theorems, write couplets, etc. But the reason will not be the programmed nature of the subject's mental life (whatever it amounts to), but will be the nature of the subject who (which) simply lacks certain sorts of programming in the first place.

On the other hand, if the subject is a certain sort of being to begin with such that certain kinds of directives, commands, and controls are possible, then the simple fact that he (it) was programmed to do this or that will not be sufficient to establish that the subject did not really do this or that.

In the case of programmed *behaviour* and those mental aspects closely associated with it the defence of La Mettrie's watch fails because with few exceptions there is no (clear or necessary) opposition between behaviour and being programmed. Indeed, behaviour seems particularly amenable to being construed in terms of well-defined tasks, and because of this it is compatible with being programmed. In their *Plans and the Structure of Behavior*, for example, Miller, Galanter, and Pribram write:

Any complete description of behavior should be adequate to serve as a set of instructions, that is, it should have the characteristics of a plan that could guide the action described. When we speak of a Plan in these pages, however, the term

will refer to a *hierarchy* of instructions, and the capitalization will indicate that this special interpretation is intended. *A Plan is any hierarchial process in the organism that can control the order in which a sequence of operations is to be performed.*

A plan is, for an organism, essentially the same as a program for a computer, especially if the program has the sort of hierarchical character described above.¹

Much current research into computer simulation of problem-solving rests on this intuition. Behaviour in so far as it can be construed in terms of well-defined tasks harbours an inbuilt compatibility with being programmed (call it 'programme-receptiveness'). And if a subject fails to possess the basic capacities necessary to the construction of routines for proving theorems, writing couplets, etc., *then* the nature (or basic capacities or input potentials) of the subject determines that the subject does not do these things and pre-empts and precludes the possibility of the subject failing in these respects because he (it) was merely programmed to effect them.

Most simply put, and contrary to the defence of La Mettrie's watch, there is *no* type of mental predicate which fails to apply to a subject simply *because* it was programmable in certain ways. Although various quasi-Rylean, behaviouristically oriented, problem-solving psychological predicates may be shown *not* to apply to a given subject, this will be because the subject lacks the capacity for being programmed in certain ways. It will not be because the subject *was* programmed in certain ways.

And when we turn to those aspects of mentality which are quite different from solving problems, playing checkers, or writing couplets—such things as having pains, emotions, after-images, etc. . . . the defence of La Mettrie's watch may be seen to be wholly irrelevant. For the having of pains, emotions, after-images, etc., are all examples of non-problem-solving non-behaviour. They are not potentially well-defined tasks which hence may be programmable, for they are not tasks at all. Consequently they could never be shown to be absent from an unmasked robot *because* the robot was merely programmed to do this or that. For they are not the sorts of things which a robot could do. They may be had, but not done.

Whether or not a robot has a capacity for pain, emotion, or after-images, will of course greatly effect what the robot could *then* be programmed to do. The problem of determining the nature of the subject is not primarily a problem of deciding how to programme it. It is a problem of knowing and deciding how to construct the subject so that it *can* be programmed in certain ways.

¹ George A. Miller, Eugene Galanter, and Karl H. Pribram, *Plans and the Structure of Behavior* (New York) 1960, p. 16.

Thus we see that the robot without a programme need not be a robot without a mind. For it could be a robot with half a mind—a robot with certain capacities for sentience, with certain input potentials, etc. It would by nature be a non-sapient robot . . . sub-moronic when compared to others of the same model which had also been equipped by benevolent manufacturers with programmes which use their basic capacities.

On the basis of the preceding remarks, then, it may be useful to sort psychological features into two kinds: those which it at least makes sense to attempt to simulate with a machine or robot through programming (in the sense of devising routines), and those which it does not. The first I shall call *programme-receptive* features of mentality. These include most kinds of problem-solving: game-playing, theorem-proving, etc. . . . i.e. rule governed activities. The second set of features I shall call *programme-resistant*.¹ These include such things as having pains, feelings, emotions, etc. It would be a methodological howler to attempt to simulate these with a machine simply by expanding current programming techniques in the sense of defining new routines. It may of course prove possible to produce machine analogues of these features through a development in hardware. (There are undoubtedly mental aspects in between those which are clearly *programme-receptive* and those which are *programme-resistant*—e.g. thinking. But I shall not go into that here.) Thus some psychological predicates may be viewed primarily as software predicates; some primarily as hardware predicates. Though I shall not develop the point, I believe this distinction may help to clarify the difference between sapience and sentience.

It is unfortunate that this rather elementary but crucial distinction is very often overlooked by both philosophers and those working on artificial intelligence and computer simulation of cognitive processes. This has resulted in philosophers coupling the mental limitation of machines (period) with the limitations as to what a machine could be programmed to do. And it has resulted in some computer simulation theorists trying to simulate programme-resistant features of mentality using simulation techniques which are strictly relevant only to the simulation of programme-receptive features: for example, there has been an attempt to extend programming techniques used in simulating problem solving behaviour so that merely by developing new routines a machine will also simulate human emotions—the having of fears, anxiety, and so on.

¹ Cf. Keith Gunderson, 'Some Mental Limitations of Some Machines', chapter IV of *Mentality and Machines* (in preparation) where the distinction between *programme-receptive* and *programme-resistant* mentalistic phenomena is worked out in detail in connection with an IPL-V computer.

To show that such a machine was only programmed to do such-and-such could not possibly show the absence of consciousness in that machine. Conversely, however, one could not make such a machine conscious simply by programming it in a certain way. Instead the link between consciousness in machines and what machines can be programmed to do is this: if a machine could be programmed in certain ways, this might demand consciousness in the machine: and if it were not possible to programme the machine in certain ways, this might suggest that the machine lacked capacities which we can loosely refer to as its consciousness.

IV

The above should help us to see why Ziff's remarks about how a robot's actions in a particular context depend on how it was programmed to act have little bearing on the question 'Could robots have feelings?' Looked at from one angle it seems that Ziff's imagined robots in the last analysis are not capable of feeling anything. But if so, the only reason they are not, is that they are specified in a vague way at the outset of Ziff's article and end up, as it were, 'retroanalytically defined' as robots which are without feelings. That is to say, at the end of his analysis we are told what his robots, about which so many interesting questions were raised at the outset, were really like. But what I have tried to show is that other 'retroanalytic definitions' are compatible with the robots so flexibly characterised at the beginning of Ziff's article.¹ And furthermore and mainly what I wish to object to is the suggestion that Ziff's robots, however interpreted, cannot feel *because* they only do what they are programmed to do, or *because* they could be programmed to respond in all sorts of strange ways to various stimuli. Ziff's programmed robots cannot feel simply because of the sorts of robots Ziff decided to say that they are; not unsimply because of the sorts of programmes they have. From another angle, however, Ziff's view becomes virtually impossible to understand. This is so if one attributes to him the claim attributed to him by Putnam where he (Putnam) says: 'Ziff has informed me that by a "robot" he did not have in mind a "learning machine" of the kind envisaged by Smart, and he would agree that the considerations brought forward in his paper would not necessarily apply to such a machine (if it can properly be classed as a "machine" at all)'.² Given this interpretation it is difficult to understand how Ziff's robots could be 'virtually indistinguishable from men in practically all respects: in

¹ Compare this criticism of Ziff with a criticism of Ayer's 'The Concept of a Person' made by Raziel Abelson in his 'Persons, P-Predicates, and Robots', *American Philosophical Quarterly* 3 (1966), 308.

² Hilary Putnam, 'Minds and Machines', in *Dimensions of Mind*, ed. by Sidney Hook, p. 176.

appearance, in movement, in the utterances they utter, and so forth'. Ziff wishes to have it both ways if he does not want his robots to be as sophisticated as 'learning machines', and yet seem as sophisticated as men. Even to *seem* that sophisticated in mental capacities strongly suggests subjects psychologically less vacant than Ziff supposes his robots to be.

Similarly, Cohen's remarks on artificial minds are mistaken in so far as he claims an opposition between *mentality* and the concept of total subservience to known or knowable rules. All that Cohen's case of the totally programmed wife shows is that someone's mind could be controlled in a certain way: it does not show that a wholly programmed mind is no mind at all. To show that someone does not think for himself, is not in the least to show that someone does not really think, and in the same sense(s) as the present sense(s) of 'think', whatever that (or they) may come to.

Arguments to the effect that a robot which only did what it was programmed to do could not be a robot which *really* was conscious, had feelings, thoughts and intentions, or which *really* deliberated and decided, fail because of the foregoing. But here let me anticipate some objections to my objections. One reason why the notion of 'being programmed' has come to seem opposed to the notion of 'being conscious' is because most of the currently existing programmed devices (computers, *et al.*) have appeared—at least to some philosophers—obviously non-conscious, so in fairness to part of the 'spirit' of some of the articles being criticised, I should add that there is some textual justification for pointing out that the authors of the articles cited are not always arguing simply from the fact that a subject is programmed to the conclusion that the subject is therefore non-conscious. They are also to some extent presupposing that there are other relevant facts which one discovers when one discovers that the subject is programmed which count against saying that the subject could be conscious. Nonetheless: (a) throughout the literature these presuppositions are never clearly set out, defended, or developed; (b) no distinction is drawn between the very different sorts of subjects which could be programmed; (c) there is no reason for supposing that for every subject which is programmed one will be able to uncover independent facts which would tend to establish that the subject was not conscious, capable of thought, creativeness, and the like. And (d) in some cases, given that the subject could be programmed to do certain things, what we would expect to find would be just the opposite: namely that the subject *was* conscious, capable of thought, intentional behaviour, and so on. (Exactly what it would take to show that such expectations were satisfied is, of course, no easy matter to decide. The 'other minds' problem could arise in connection with robots—even though we made them.)

v

It may be argued that though it would be possible to programme a robot to have or effect all sorts of marvellous (humanoid) experiences and accomplishments if it had certain basic capacities to begin with, it is an analytic truth that robots could never possess the requisite capacities for such experiences or accomplishments. But those who would argue that a subject could not be both a robot and have certain basic capacities or input potentials, may be in the same position (now) as James and others were in towards the end of the 19th century when it was vehemently maintained that a subject could not be both wholly mechanistic and exhibit self-adaptive behaviour. This also had the flavour of an analytic truth before various negative feedback machines were developed and the notion of 'self-adaptive behaviour' received further clarification.¹

If psychological predicates are really applicable only to persons, which I think is unlikely, then robots of certain sorts might be persons of certain sorts. They would be persons to which certain other hitherto non-person predicates applied—certain clusters of what might be called B-(biographical) and E-(ecological) predicates such as, 'is manufactured', 'was a 1966 model', 'sleeps electrically', etc.

There is, however, no more reason to assume that robots would have to be persons in order to partake of psychological predicates than there is reason to assume that a dog would have to be a person to be able to show affection and obeisance. In other words, I believe that psychological predicates can become applicable to machines or robots, as I think some already have, without blurring the line between robots and persons. In short it seems possible that such a line need not be blurred *except with respect to the predicates in question*. That is, it may be possible to show that a machine could, for example, recognise various sorts of patterns without blurring the line between men and machines *except in so far as recognition is concerned*. If in such a case the line between men and machines need not be blurred except with respect to 'recognition' (and its near synonyms) then it is hardly clear that the meaning of such words as 'men', 'persons', 'robots', or 'machines', will have altered. Both rolling stones and rolling men gather no moss, which hardly shows that men are borderline cases of stones. When Thomas Edison invented the light bulb he did not produce a borderline case of a white dwarf star. (Compare 'If it really throws light it's some sort of star' with 'If it really has feelings it's some sort of living organism'.)

In the case of living things and machines there was no global blurring of

¹ Cf. Keith Gunderson 'Cybernetics', in *Encyclopaedia of Philosophy*, edited by Paul Edwards (Crowell-Collier, 1967).

the line between men and machines when we began (appropriately) to refer to 'self-adapting' machines. Instead a localised sharing occurred in connection with the feature of self-adaptation. That is precisely why cybernetics seemed to provide a non-trivial counter example to some of the claims made by vitalists.

So to build robots to which all sorts of psychological predicates legitimately apply need not be comparable to building robots which were just like us. They would of course be rather like us in so far as they share those predicates. And to be like us with respect to sharing a given capacity or the exercise thereof does not necessitate that the robot be identical to us even in that respect. Compare both washing machines and human beings wash clothes; but they are not identical with respect to this capacity or the exercise thereof. Hence for a robot to have psychological predicates apply to it is not automatically, as many would assume, either the same as the robot being the same as, or even being almost the same as, a person.

Furthermore, it is not even the case that in every instance where the concepts were blurred, interesting discoveries would be ruled out. For how the concepts (of men and machines) come to get blurred is also of central importance. If, for example, we found that we do not really do some things which we thought we did, and are therefore more like certain machines than we had hitherto realised, or if we found out that the way that we did things was more purely mechanised than hitherto imagined, then in both cases something very interesting might be found out. (Compare recent suggestive attempts by Newell, Shaw, and Simon to explain creative thinking in human beings in terms of certain kinds of rule-governed behaviour.)¹

A view more explicitly of the kind the immediately foregoing remarks contrast with is developed by A. R. Lacey in his article 'Men and Robots' where he writes:

Let me add again that I do not mean that no artifact will ever achieve feeling, but that such an artifact would not be a robot, in my sense, but a human, or on the human side of the fence.²

But it is not at all certain that in every case where artifacts had feelings they should for that reason be placed 'on the human side of the fence' (except trivially, for example, with respect to feelings). And I disagree with him where, after despairing of finding the differences between men and robots in terms of what they can do, he writes:

This does not mean that we could not create an artifact which turned out to be conscious; but such an artifact would be a human (or some such thing) and not a robot in our present sense.³

¹ 'The Processes of Creative Thinking', the RAND Corporation, p-1320, 16 September, 1958.

² *Philosophical Quarterly*, 10 (1960), 61-72.

³ *Ibid.*

His qualifying phrase 'in our present sense' does not barricade Lacey's claims against the linguistic possibility that 'conscious' could acquire new uses (without altering its meaning)¹ in much the same manner that 'problem solving' and 'self-adaptive' seem to have done. It remains possible in the face of such diachronic surprises to trivialise a thesis like Lacey's by, say, simply not identifying the present sense of 'robot' with the then current senses of 'robot'. One can always choose to use a word in an idiosyncratic manner, but nothing is gained thereby. (There is no such thing as an idiosyncratic *sense* of a word.) And obviously Lacey is not doing that. So if he does wish to identify his construal of the meaning of the word 'robot' with *the* meaning of the word 'robot', his claims based in part on this identification will be vulnerable.

There need not be any 'the difference' between men and robots. Different differences are possible and must be discussed in any attempt to stake out what a robot could be and do and how what it could be or do would serve to distinguish men from robots. Some of these differences would no doubt be marked out by the use of psychological predicates and some would be marked out by the use of what I have loosely called B- and E-predicates. Some by others.

Many philosophers including many defenders of La Mettrie's watch have looked for an almost magical covering word or predicate (or small set of predicates) such that if it or they (always) applied or failed to apply to a certain subject we would then *be safe* (always) in saying that the subject was definitely (or definitely not) a human, or a robot, or a machine. And so they persist in perpetrating such claims as 'If a subject could think (create, feel, perform, purposively, and so on) then it would not be a robot (or machine, etc.)' and have failed to note that no one feature (or small set of features) such as 'feels' or 'is made in a factory' can be regarded as essential to something being or not being a robot or human being or machine. Perhaps a programmed conscious robot could perform as well as a human being, and do this by, for example, thinking out the answer, and still not thereby be a kind of us. Many other relevant differences might remain: biographical, ecological, and so on.

VI

True borderline-case decisions may be as rare in philosophy as true tied-finishes are in the Kentucky Derby. So far, at least, the question of whether an actual subject is a humanoid machine or a mechanistic human has not

¹ Cf. R. Puccetti's interesting article 'On Thinking Machines and Feeling Machines', this *Journal*, 18 (1967), where a parallel use is made of Putnam's important distinction between (1) a word acquiring a new *use* in the language because of the (core) meaning it has (had) and (2) a word being (arbitrarily) *given* a new use (or meaning).

arisen. Instead the question has arisen as to whether certain predicates ('recognises patterns', 'plays chess', 'writes poetry', etc.) hitherto used only in connection with human beings should apply to what are most certainly machines; or whether certain predicates hitherto used only in connection with machines (e.g. 'is programmed') also apply to what are most certainly human beings.

In a recent article, 'Robots: Machines or Artificially Created Life'¹ Hilary Putnam seems to me to mischaracterise the central issues involved in 'minds and machines' controversies in so far as he assimilates the question of clarity 'with respect to the "central area" of talk about feelings, thoughts, consciousness, life, etc.' to the question of clarity 'with respect to the "borderline-case" of robots (i.e. as machines or artificial life)'. If and when we draw open the closed curtain of technological ignorance and find persots or robsons inhabiting the parlours of persons, we will probably (though not necessarily)² know what to say about them. If problems persist which are to be solved by making a linguistic decision, they will be trivial problems: for in a clear case of a border-line case the only decision involved in selecting descriptions is a poetic one (e.g. shall we call them 'persots' or 'robsons'? I myself prefer 'persots'). To think otherwise is to mistake neology for ontology. A tangerine is a tangerine is a tangerine.

The non-trivial, non-science-fictional problems of philosophy now are rather such problems as: Are the information processes underlying the problem-solving capacities of, for example, a (clearly non-human) list-processing computer analogous to human ones? Do so-called pattern-recognition programmes actually enable machines to *recognise* patterns? What sorts of evaluation procedures can be developed which will assist us in deciding which of any two competing simulations are the most useful in speculating about human mentality. Which machine models of human mentality are subject to plausible neurophysiological interpretation, which ones are not? What relationships obtain amongst mentalistic phenomena which are receptive to being simulated by current computer programming techniques, and those which are not? Which psychological predicates refer to basic capacities, and which apply to behaviour 'generated' by those capacities? Which are a mixture? Short of making a mind, how can we model aspects of the mind?³

University of Minnesota.

¹ *J. Phil.*, 61 (1964), 668-91.

² Cf. my 'Cybernetics and Mind-Body Problems', *op. cit.*

³ Cf. Keith Gunderson, 'Philosophical Considerations of Artificial Intelligence and Computer Simulation of Cognitive Processes', chapter III of *Mentality and Machines* (in preparation) where all these questions are discussed.

Corroboration, Explanation, Evolving Probability, Simplicity and a Sharpened Razor

by I. J. GOOD

I. INTRODUCTION AND SUMMARY

My aim is to discuss (degrees of) corroboration, explanatory power, and complexity, and the relationships between them, and to suggest a new form of Ockham's razor which is applicable in particular to practical statistical problems. It will be found necessary to distinguish between two meanings of 'explanatory power' but nearly unique explicata will be suggested in terms of probability. These explicata will be derived from reasonable desiderata or axioms. Corroboration, explanatory power, and complexity can be expressed in terms of one another, and probability can be defined in terms of them, but I believe that the best practical method for estimating any of these things is to make judgments about all of them, and perhaps other things as well, and then to make deductions and modify the judgments in the light of inconsistencies that emerge.

Among the conclusions are (a) (degree of) corroboration can be identified with 'weight of evidence' (in a sense not to be confused with the sense in which it was used, for example, by Popper, 1959), where the identification preserves all ordering relations (inequalities) between degrees of corroboration; (b) the explanatory power in the 'weak sense' of a hypothesis H , for explaining evidence E , namely the explanatory power that is unaffected by cluttering up H with irrelevancies, can be identified with the amount of mutual information between H and E ; (c) the explanatory power in the strong sense, that is affected by the cluttering mentioned, can be obtained by subtracting from the weak explanatory power a term proportional to the information in H ; (d) the complexity of a proposition can be identified with the amount of information in it, that is, minus the logarithm of its probability; (e) a sharpened form of 'Ockham's razor' is proposed which is that if our primary purpose is explanation we should select the hypothesis (among those we know) which has the maximum strong explanatory power; (f) this improvement of Ockham's razor can be applied to the estimation of statistical parameters. It will typically give rise to interval estimates, but it would give point estimates if weak explanatory power

Received 27.xii.67

were used. It can also give point estimates, in a sense, if non-zero probabilities are associated with certain parameter values: this would correspond to the acceptance of null hypotheses.

We shall usually assume that hypotheses have probabilities: this assumption is the real essence of the Bayesian approach to statistics. The probabilities of hypotheses are nearly always subjective in practice, although they can be regarded as judgments of logical probabilities. Nevertheless it turns out that corroboration can sometimes be objectively measured, and our main new desideratum for corroboration is in brief a demand for objectivity in some circumstances. This desideratum was not assumed in Good (1960b), which continued the work of Popper (1959), but two of the less compelling desiderata assumed before will not be assumed here. Instead they can be inferred as theorems. In the previous work, the explicatum for corroboration was not unique. Also the double meaning of 'explanatory power' has previously been overlooked.

The objectivity of corroboration can be extended by regarding it not as a scalar but as a function of one or two variables. This remark will be explained in Section 4(c).

We shall be forced to the view that explanatory power is nearly always subjective, and is even usually based on 'superficial' or 'evolving' probabilities, the kind that must be used when referring to the probability that a mathematical proposition is true, or to the information provided by a calculation. Evolving probabilities are required in order to refute some of Popper's views on simplicity.

We shall omit some mathematical proofs: where required they are mostly explicit or implicit in Good (1960b and 1966b).

2. NOTATION

Let E usually represent (the proposition expressing) some of the evidence or events obtained from some experiments or observations. (Typically some of the observations are judged to be irrelevant and are not reported. We can think of E as a *class* of results each of which would have been described in the same way. For example, when we have tossed a penny we might describe the result as 'heads' without specifying that the Queen's chin was pointing South.) Let H usually represent a hypothesis for which $P(E|H)$, the probability of E given H , has a known or judged value for each E that might have been reported. Here H might denote a physical theory or hypothesis, or a statistical hypothesis. If it is a 'simple' statistical hypothesis, then $P(E|H)$ has, for each E of some specified set, a numerical value known as part of the *definition* of H , so that in this case $P(E|H)$ is

a 'tautological' probability. Here the use of the word 'simple' is technical and is not to be confused with the ordinary notion of simplicity which we shall also discuss. The relative magnitudes of $P(E|H)$, for various H 's, are known as the *likelihoods* of the hypotheses. If H is a *composite* statistical hypothesis, then $P(E|H)$ has a definite value only if the relative initial probabilities are known or assumed for the simple statistical hypotheses of which H is the logical disjunction. [Partly in order to avoid the integral calculus I often assume that none of these probabilities is zero. This is possible since the number of possible hypotheses is at most countably infinite: even the parameter values can in practice take only 'computable' values in the sense of Turing, provided that they are specified in a finite number of words and symbols: continuity is introduced only for mathematical convenience. See Good (1950), p. 55.] These initial probabilities are nearly always subjective (personal) or multisubjective so that $P(E)$ is then also subjective, where $P(E)$ is an abbreviation for $P(E|G)$ and G is taken for granted. (I shall subsume 'multisubjective' under 'subjective'.) Subjective probabilities are assumed to be based on mature judgments and the 'subject' (known as 'you') is supposed to have tried fairly hard to make them consistent with the axioms of probability. But owing to the complexity of the world, of science, of business, of chess, you are usually forced to make your judgments without complete logical analysis. This is invariably the case when judging the probability that a mathematical theorem is true, or that a chess game would be won with best play, unless the probability reaches 0 or 1. The probabilities change according to how far you have carried the analysis and we may reasonably call them *shifting, slippery, sliding, superficial, surface, changeable*, or *evolving probabilities*. (See Good, 1950, p. 49 and 1967b, pp. 92, 108, 114.) The axiom that $P(E|H) = P(F|H)$, if E and F are logically equivalent given H (that is $E.H$ implies F and $F.H$ implies E), has to be replaced by $P(E|H) = P(F|H)$ *when you have proved that E and F are equivalent given H* . Evolving probabilities cannot be logically consistent within classical logic: in practice you can hope to be at best reasonably consistent so perhaps nearly all subjective probabilities evolve: it is a matter of degree. (See also Good, 1962, for the relationship of this with two kinds of rationality.) An evolving probability $P_t(E|H)$ depends on the time t at which it is judged, although in this paper I shall usually omit the subscript t for the sake of simplicity of notation. This philosophy should appeal to the mathematical intuitionist. Compare also the approach of Poincaré (1913/1963, Chap. 4) to the antinomies of logic.

If H is a physical theory then $P(E|H)$ might be tautological, physical, subjective, evolving (and subjective), or (rarely) logical. (For a discussion of kinds of probability, other than evolving ones, see Good, 1959a, 1966a.)



We shall also use the notation $P(E|H)$ where E and H are propositions that do not necessarily refer to events and hypotheses. Propositions will be denoted by E, F, G, H, K , sometimes with subscripts.

Let $O(E|G) = P(E|G)/P(\bar{E}|G)$, the odds of E given G , where the vinculum denotes negation. Similarly, if E and F are mutually exclusive, let $O(E/F|G) = P(E|G)/P(F|G)$, the odds of E as against F , given G , also expressible as $O(E|G.(E \vee F))$, where $E \vee F$ denotes the disjunction of E and F . Let $I(H:E|G) = \log \{P(E|H.G)/P(E|G)\}$, the information concerning H provided by E given G . It is also called the 'mutual information' between H and E given G since it is symmetrical in E and H when $P(E.H|G) \neq 0$. Here $H.G$ denotes the logical conjunction of H and G . The colon denotes 'provided by' and must be distinguished from the vertical stroke. The base of the logarithms can be any real number greater than unity; its choice is unimportant in the present context since it merely determines the unit in terms of which information (or weight of evidence: see below) is measured. Information is the logarithm of the 'association factor', $P(E.H|G)/(P(E|G)P(H|G))$. Several justifications for calling I 'information' have been published: a justification making minimal assumptions is given by Good (1966b). In particular $I(H:H|G) = -\log P(H|G)$ is abbreviated to $I(H|G)$ and is called the amount of (semantic) information in H given G . Amounts of semantic information are logical, subjective, or evolving depending on the kind of probabilities (Good, 1950, p. 75; and, in much detail, Bar-Hillel and Carnap, 1953). Let

$$\begin{aligned} W(H:E|G) &= \log \{P(E|H.G)/P(E|\bar{H}.G)\} \\ &= \log \{O(H|E.G)/O(H|G)\} \\ &= I(H:E|G) - I(\bar{H}:E|G), \end{aligned}$$

the weight of evidence concerning H provided by E given G . Weight of evidence is the logarithm of the (Bayes-Jeffreys-Turing) 'factor in favour of a hypothesis', the factor by which the initial odds are multiplied to get the final odds. Sometimes G is taken for granted and omitted from the notation. *We are liable to omit G without warning.* For justification of the name 'weight of evidence' see, for example, Good (1950, 1960-1), where references to H. Jeffreys and A. M. Turing are given. (When H and $\bar{H}$ are both simple statistical hypotheses, weight of evidence is identical with the logarithm of a 'likelihood ratio', but not more generally since 'likelihood ratio' is currently given an *objective* and important meaning, namely the ratio of *maximum* likelihoods, a concept introduced by Neyman and Pearson, 1928.) The reason W was introduced into statistics was that it seemed to measure weight of evidence, corroboration, confirmation, or support, in the *English* senses of these expressions.

It will be noted that we have referred to the probability of E given $\bar{H}$. As before, this is a tautological probability if $\bar{H}$ is a simple statistical hypothesis, otherwise it will usually be a subjective probability, or in particular an evolving one.

Let $C(H:E|G)$ denote the degree of corroboration of H and let $D(H:E|G)$ denote the weak-sense (or just 'weak') explanatory power of H , for explaining E , given G , or provided by E given G . Our main aims at first are to find explicata for $C(H:E|G)$ and for $D(H:E|G)$. We at first think of $C(H:E|G)$ as a single (scalar) number, but we shall ultimately argue that it depends on what alternatives to H we have in mind, and for each alternative H_1 we have a different corroboration $C(H/H_1:E|G)$. These various corroborations can be collapsed into a single scalar if the values of the initial probabilities of the various simple statistical hypotheses are assumed.

3. THE DESIDERATA AND EXPLICATA OF EXPLANATORY POWER AND CORROBORATION IN TERMS OF PROBABILITIES

Any real function of $P(E|H)$ and $P(E|\bar{H})$ that, together with $P(H)$, mathematically determines $P(H|E)$, is a monotonic function of $W(H:E)$. (A fairly complete outline of a proof of this is given in Appendix A.) This might by itself be regarded as a fair enough justification for taking a monotonic function of $W(H:E)$ as an explicatum for $C(H:E)$, but the following desiderata are I think more compelling. We call the desiderata 'axioms'. Axioms 1 to 10 and Theorems 1, 14, and 18 are the same as the corresponding axioms and theorems in Good (1960b) except that Axioms 8 to 10 assume less than before. Axioms 1 to 10, in the present version, assert so little that, for them, little discussion of the meanings of 'corroboration' and 'weak-sense explanatory power' seems necessary. It is necessary only to emphasise the distinction between explanatory powers in the weak and strong senses. As we said in the Introduction the former is to be unaffected if H is cluttered up with irrelevancies. This is expressed by Axiom 10. For strong explanatory power, denoted by D' , we assume that D can be replaced by D' in Axioms 1, 3, 4, 7, 8, and 9, but not in Axioms 2, 5, 6, and 10. Axiom 10 is clearly not wanted for D' , Axioms 2 and 5 are not at all compelling for it, and Axiom 6 is irrelevant when Axiom 5 is not assumed.

Axiom 1. $C(H:E|G)$ and $D(H:E|G)$ are continuous differentiable real functions (taking values that might be positive, negative, or zero) of the 'absolute' and conditional probabilities of logical combinations of H , E , and G , where the permitted logical operations are conjunction, disjunction, and negation. (There are 52,670 of these probabilities that are not necessarily equal to 0 or 1.) An 'absolute' probability is one in which the 'given' proposition is taken for granted. References to differentiability, and later to

analyticity, are probably unnecessary, as J. Aczél has pointed out (private communication), but the proofs are easier if they are assumed.

Theorem 1. $C(H:E|G)$ and $D(H:E|G)$ are continuous differentiable real functions of $P(H)$, $P(E)$, $P(G)$; $P(E.G)$, $P(G.H)$, $P(H.E)$, and $P(H.E.G)$. (Theorem 1 is a trivial consequence of Axiom 1 and conversely.)

Theorem 1A. $C(H:E)$ and $D(H:E)$ are continuous differentiable functions of $P(H)$, $P(E)$, and $P(H.E)$ and therefore of $P(E)$, $P(E|H)$, and $P(E|\bar{H})$. (Theorem 1A, and Theorem 1B below, are again trivially proved.)

Axiom 7. (Relativisation.) Whatever functions $C(H:E|G)$ and $D(H:E|G)$ are of the various absolute and conditional probabilities then $C(H:E|G.K)$ and $D(H:E|G.K)$ are the same functions of these probabilities made conditional on K .

Theorem 1B. $C(H:E|G)$ and $D(H:E|G)$ are continuous functions of $P(E|G)$, $P(E|H.G)$, and $P(E|\bar{H}.G)$. We denote these functions by $\phi\{P(E|G), P(E|H.G), P(E|\bar{H}.G)\}$ and $\psi\{P(E|G), P(E|H.G), P(E|\bar{H}.G)\}$.

Axiom 2. $C(H:\bar{H}|G) \leq C(H:E|G) \leq C(H:H|G)$, and similarly for D .

Axiom 3. $C(H:\bar{H}|G) < C(H:H|G)$ except perhaps when $P(H|G) = 0$ or 1 . Similarly for D .

Axiom 4. If $C(H:\bar{H}|G) < C(H:H|G)$ and $C(H:E|G) = x$, then F exists such that $C(H:F|G)$ takes any assigned value sufficiently close to x . Similarly for D .

Axiom 5. If evidence is considered in two parts, E and F , then the corroboration of H is determined by that provided by E , combined with that provided by F when E is known. More precisely, there is a fixed continuous analytic function f such that

$$C(H:E.F|G) = f\{C(H:E|G), C(H:F|E.G)\},$$

provided that E and F do not conflict, in the sense that

$$P(H|E.G) = 0, P(H|F.G) = 1$$

or

$$P(H|E.G) = 1, P(H|F.G) = 0.$$

Similarly for D in place of C .

Let a and b be the least and greatest possible values (strictly the least upper bound and greatest lower bound) of degree of corroboration, where a and b might be finite or infinite.

Axiom 6. If $a < x < b$, and $a < y < b$, then $f(x, y)$ is strictly increasing in each variable when the other is held constant. Also $f(a, y)$, $f(b, y)$, $f(x, a)$, and $f(x, b)$ are non-decreasing functions of y and x respectively.

Axioms 8, 9, and 10 will be assumed for D only, not for C . Otherwise they are the same as in Good (1960b).

Axiom 8. $D(H:E|H) = 0$ if E does not imply $\bar{H}$. In words, if H is already assumed, to say it again does not help to explain E any further.

Axiom 9. If $D(H:E|G)$ depends only on $P(H|E.G)$ and $P(H|G)$, then it is a strictly increasing function of $P(H|E.G)^*$ when $P(H|G)$ is fixed.

Axiom 10. Let K be 'nothing to do' with H and E , given G ; more precisely suppose that K is independent of H given G , is independent of H given $E.G$, and is independent of E given G . Suppose also that $P(K|G) > 0$. Then $D(H:E|G) = D(H.K:E|G)$.

Theorem 14. $I(H:E|G)$, and strictly increasing differentiable functions of it, form the only explicata for $D(H:E|G)$. (The proof of this theorem, which is not trivial, is given in Good, 1960b.)

Thus, weak-sense explanatory power can be identified with information while preserving all inequalities, and we shall make this identification by convention.

Note that the function ψ , of Theorem 1B, depends only on its first two arguments. If this had been assumed we could have inferred Theorem 14 from Axioms 2 to 6 alone. (See Good, 1960b, Theorem 17.)

Popper (1959) has argued, surprisingly, that a simpler theory is (initially) less probable than a complicated one. This seems to me a misleading usage of terminology and I think Popper should recant since $H.K$ is more complicated and less probable than H when H and K are mutually irrelevant. But consider the following argument on Popper's side. (It was brought to my attention, in support of Popper's view, by J. Agassi, about 1960, except that he referred to Fresnel's optical laws and Maxwell's equations in place of planetary orbits and the inverse square law.)

The inverse square law H approximately implies the planetary motions E . Therefore $P(E) \geq P(H)$. But E seems much more complicated than H and Popper might seem to be vindicated.

My reply to this is that $P(E)$ suddenly increased greatly, as judged by the astronomers of the time, as a consequence of Kepler's and Newton's calculations. I do not mean merely that $P(E|H) > P(E)$, but that $P(E)$, now judged to satisfy $P(E) = P(H) P(E|H) + P(\bar{H}) P(E|\bar{H}) > P(H) P(E|H) \approx P(H)$, is judged to be much larger, in ratio, than it had been judged to be before it was noticed that $P(E|H) \approx 1$. The motions of the planets merely *seemed* to be complicated, and were simplified by pure thought. We are *forced* to the view that explanations depend on *evolving* probabilities when the explanation does not involve any new empirical observations. It is surprising in retrospect that this argument was ever overlooked.

We are now in a position to discuss explanatory power in the strong sense. We denote it by $D'(H:E|G)$ and have already assumed that it satisfies Axioms 1, 3, 4, 7, 8 and 9, with C replaced by D' although the following

argument makes use only of Axioms 1, 3, 4, and 7. I have not worked out quite as much detail as for C and D and so the desiderata will not be called 'axioms', but the following assumptions are formally extremely reasonable and the explicata reached are the only ones possible. As often, G will be taken for granted.

The first assumption is that $D'(H:E)$ is either a function of $P(H)$ and $D(H:E) = I(H:E)$ or, what is easily seen to come to the same thing, a function of $P(H|E)$ and $D(H:E)$. Next we assume that, if H and E have nothing to do with K and F , then $D'(H.K:E.F)$ is some function of $D'(H:E)$ and $D'(K:F)$. In accordance with Appendix B, these assumptions lead, up to a differentiable monotonic transformation, to the general solution:

$$D'(H:E) = I(H:E) - \gamma I(H) = (1 - \gamma)I(H:E) - \gamma I(H|E)$$

(and the function we just assumed to exist becomes addition). If H is an obvious tautology (or is obviously deducible from G), D' reduces to zero as it should, since an obvious tautology explains nothing new. If H is E then D' becomes proportional to $I(E)$, whereas if H has nothing to do with E then D' becomes proportional to $-I(H)$. We must assume $0 < \gamma < 1$ and the value of γ depends on the relative weights we wish to give to explanation and to the avoidance of 'clutter'. We could give equal weights to them by taking $\gamma = \frac{1}{2}$ and this gives the simplest explicatum for D' . Taking $\gamma = 0$ gives the weak-sense explicatum which can therefore be regarded as a limiting case of the strong one, and $\gamma = 1$ would assign maximum explanatory power to the simplest possible hypothesis, namely an obvious tautology. Thus it is absurd to take $\gamma = 1$, but if a tautology gives as large a value to the evolving probability $P(E|H)$ as any other hypothesis does, then indeed the tautology would provide the best strong explanation of E . This would be so if we could show, in Kapp's phrase (c.1961), that 'Nature's Statute Book is empty', that the universe is the only possible one. Such a tautology is not an obvious one to say the least.

So our explicatum is

$$\begin{aligned} D'(H:E|G) &= I(H:E|G) - \gamma I(H|G) \quad (0 < \gamma < 1) \\ &= \log \{P(E|H.G)[P(H|G)]^\gamma / P(E|G)\}, \end{aligned}$$

and $\frac{1}{2}$ is a sensible value for γ . When $\gamma = 1$ we get simply $\log P(H|E.G)$ which as we said is too extreme. The constant γ measures the degree to which the simplicity of the hypothesis is regarded as desirable (since $I(H|G)$ will later be identified with the complexity of H) as compared with its weak explanatory power. We thus have a continuum of strong explanatory powers.

If H implies E , for example if H is E itself, you are still interested in explanations for E unless H is certain. A clear statement of E is a full

enough explanation of the *meaning* of E and is a full explanation in the weak sense, but this is only half the battle (literally so if $\gamma = \frac{1}{2}$) since ideally we want $P(E|H) = 1$ and $P(H) = 1$, the strong sense of *full* explanation, the Empty Statute Book sense.

A comparison of the explanatory powers of two hypotheses is usually of more interest than the estimation of their individual powers. We obtain natural comparisons by subtraction. We can denote the hypotheses by H and $\bar{H}$ if we take their disjunction for granted and we have (see Theorem 18 below):

$$\begin{aligned} D(H:E) - D(\bar{H}:E) &= I(H:E) - I(\bar{H}:E) = W(H:E) = C(H:E), \\ D'(H:E) - D'(\bar{H}:E) &= W(H:E) + \gamma \log O(H) \\ &= (1 - \gamma)W(H:E) + \gamma \log O(H|E). \end{aligned}$$

Thus a hypothesis is of large explanatory power, in the weak sense, in comparison with its negation, if it is well corroborated, and, in the strong sense if it scores well when the weight of evidence is added to a *fraction* of the initial log-odds. For example, if $\gamma = \frac{1}{2}$, which is the simplest value, it is somewhat as if we used Bayes-Jeffreys-Turing factors but with the initial odds replaced by their square roots.

As a simple example of D' , consider the question of why space has three dimensions (proposition E). Since 3 is a small and therefore simple integer, $P(E)$ is not very small and $-\log P(E)$, which is the maximum explanatory power for the theory, irrespective of the value of γ , is not very great. In short, there is *not very much* to explain, although explanations would be very much worthwhile because the problem is so fundamental, and a good explanation would presumably also explain other things. (How we come to *believe* that space has three dimensions is another question requiring explanation. See Poincaré 1913/1963, chap. 3; Zeeman, 1965.) An explanation would have been in much greater demand (Good, 1959b) if the number of dimensions had been 1,259,482,504 since $P(E)$ would have been small.

For reasons clarified later, we call the principle of selecting the hypothesis of greatest explanatory power in the strong sense, the 'Sharpened Razor'. It can be applied in statistics. Usually a statistician considers classes of hypotheses indexed by one or more real parameters, such as the hypotheses that a binomial physical probability is p ($0 \leq p \leq 1$). If he is a Bayesian he will often assume an initial (prior) distribution for p , corresponding to our probabilities $P(H)$. Let us apply the Sharpened Razor in order to estimate p from a binomial sample consisting of r 'successful' trials out of n (event E). The result is to select a *set* of values of p , usually consisting either of a single value of p ('point estimation') or an interval of values ('interval

estimation'). If the initial distribution of p has a density function, then point estimation (taken at its face value) would be ruled out because $I(H) = \infty$, and so $D' = -\infty$ for each hypothesis that fixes p . (In practice point estimates are implicitly intended only as approximations but I have taken them literally.) For a binomial sample with r and $(n-r)$ both large and with a uniform initial distribution for p , I find that the hypothesis of greatest strong explanatory power is approximately defined by

$$(p-r/n)^2 \leq r(n-r)\xi/n^3,$$

where $\xi = \xi(\gamma)$ is the unique positive solution of

$$\xi = (1-\gamma)e^{\xi^2/2} \int_0^\xi e^{-x^2/2} dx. \quad [\xi(0) = 0, \xi(\frac{1}{2}) = 0.92, \xi(\frac{1}{2}) = 1.40, \xi(\frac{3}{4}) = 1.94, \xi(1) = \infty.]$$

Presumably it is typical that the linear dimensions of the set are roughly inversely proportional to the square root of the sample size, as in ordinary Bayesian statistics. When $\gamma = 0$, the razor has been sharpened too much; we have 'weak-sense' estimation, and the interval estimate is cut down to the maximum-likelihood point estimate. When $\gamma = 1$, the razor is too blunt: it fails to cut away anything from the unit interval for p .

If, following Jeffreys (1961), we were to associate positive initial probability with a special value of the parameter, corresponding to the null hypothesis, then even with $\gamma > 0$ we would often get this parameter value as our estimate, that is, we would accept the null hypothesis. This, however, is not what is usually meant by 'point estimation'.

We shall return to the discussion of explanatory power and the Sharpened Razor in our treatment of complexity in Section 5.

For the present we continue with our analysis of corroboration.

Although it seems reasonable that explanatory power should nearly always involve a large subjective element it is desirable, both in the interests of legal and scientific applications, that degrees of corroboration should be nearly objective in as wide circumstances as possible. Moreover ordinary linguistic usage seems to ascribe a reasonable degree of objectivity to corroboration in some circumstances. If we can achieve as much as this we shall gain something and lose nothing until some new compelling desideratum is proposed that conflicts with our explicatum. A reference to Theorem 1B then shows that ϕ should be taken as mathematically independent of its first argument, since this argument is nearly always subjective, whereas the other two are objective when H and $\bar{H}$ are simple statistical hypotheses. If corroboration does not have this property it will almost never be objective. [An example in physics in which H and $\bar{H}$ could be regarded as simple statistical hypotheses was the experiment concerning the Mossbauer effect relating to gravitational frequency shift (see Good, 1960a). Another

example would be the experiments on the deflection of light grazing the sun during an eclipse, but the probable errors of some of these experiments are unreliable: see McVittie (1965). As it is, subjective judgment is required for deciding whether the deviations from the predictions of General Relativity are too large, but it is obvious that General Relativity receives enormous corroboration *in comparison with Newtonian physics* from the observations. See also Section 4c below.]

Accordingly, we assume

Axiom 13. If $C(H:E|G)$ depends only on $P(H|G)$, $P(E|H.G)$, and $P(E|\bar{H}.G)$, then it depends only on $P(E|H.G)$ and $P(E|\bar{H}.G)$. This axiom was not assumed in Good (1960b) and a wider class of explicata was derived.

Theorem 18. $C(H:E|G)$ is a strictly increasing differentiable function of $W(H:E|G)$. We can therefore identify corroboration with weight of evidence by convention.

The axioms corresponding to the present Axioms 8 and 9, with C in place of D , and also Axioms 11 and 12 of Good (1960b) now follow as theorems. But if D is replaced by C in Axiom 10, and the conditions $C(H:E|G) > 0$ and $P(K|G) < 1$ are appended then it becomes false, as can be seen without much difficulty with the help of Theorem 18. This feature of Axiom 10 formed the basis of Popper's distinction between corroboration and explanatory power and was taken as a part of the axiom in Good (1960b). I now find Axiom 13 much more compelling than this assumption especially when the latter is combined with Axioms 11 and 12 of the earlier paper. Axiom 13 has the further merit of leading to an essentially unique explicatum for corroboration.

4. FURTHER DISCUSSION, MAINLY CONCERNING CORROBORATION

(a) Carnap (1950) used the term 'confirmation' for logical probability, and the term has unfortunately been used in this sense by many philosophers of science, but by no statisticians as far as I am aware. Since in ordinary English 'confirmation' and 'corroboration' are nearly synonymous, Carnap's usage is misleading and Popper (1959, p. 393) claims that it even misled Carnap. It also misled Hacking (1965): see my review, Good (1967d). Popper (1959, p. 394) says that 'I regard the doctrine that *degree of corroboration or acceptability cannot be a probability* as one of the most interesting findings of the philosophy of knowledge'. This finding, with the words 'or acceptability' definitely deleted, was taken for granted by Jeffreys (1936), Turing (1940) and Good (1950). The term 'confirmation' in the sense of 'probability' should be abandoned since it is not clearly intended to be a technical term, and if it were a technical term it would

serve no purpose other than a trade union one. If Carnap would recant on this point of terminology he would help to put philosophers on speaking terms with other people. If you had $P(H|E)$ close to unity, but less than $P(H)$, you *ought not* to say that H was confirmed by E .

(b) The convenience of the identification of D with I and C with W (see Theorems 14 and 18), rather than with fixed functions of them, arises because of various simple properties of information and weight of evidence, such as those mentioned in Section 2 and those listed in the remainder of this sub-section.

- (i) $W(H:E.F|G) = W(H:E|G) + W(H:F|E.G)$,
- (ii) $I(H:E.F|G) = I(H:E|G) + I(H:F|E.G)$
- (iii) If E and F are independent given $H.G$ and given $\bar{H}.G$, then,
 $W(H:E.F|G) = W(H:E|G) + W(H:F|G)$,
 $I(H:E.F|G) = I(H:E|G) + I(H:F|G)$.
- (iv) If evidence relevant to H is collected in a large number of statistically independent small pieces, then the distribution of the total weight of evidence has roughly a normal distribution.
- (v) 'Expected' weight of evidence for a true hypothesis is non-negative, that is
 $\mathcal{E}\{W(H:E|G)|H\} \geq 0$ (Turing, 1940; Good, 1950, p. 72).
- (vi) Expected weight of evidence is a good explicatum for 'diagnosability' in medicine and elsewhere (Good, 1967a).
- (vii) It can be used for estimating the sample size of a sequential binomial experiment (Turing, 1940) and can be used in the design of more general experiments when utilities are too difficult to estimate; see Good (1960b).
- (viii) Expected amount of information is basic in Shannon's theory of communication (Shannon and Weaver, 1949), and is related to entropy in statistical mechanics. It can be used in the design of experiments: see Good (1956) and for several detailed examples, Lindley (1956). Lindley refers to an unpublished 1953 report by L. J. Cronbach. Fisherian information is an alternative in some cases.
- (ix) Expected weight of evidence is also closely related to entropy. It can be used in order to expound and elaborate on the coding theorems of communication theory (Good and Toulmin, 1967) and in order to reformulate the 'principle of maximum entropy' (Good, 1966a; Kullback, 1959). It can also be used in an explication of 'decision' (Good, 1964).
- (x) Weight of evidence is also relevant to the explication of probabilistic causality, in the writer's opinion (Good, 1961).

It is sometimes convenient to leave out the logarithms; see, for example, the theorem of the weighted average of factors in Good (1950, pp. 68, 91). The use of the antilogarithms also eliminates the need to refer to a unit of measurement of evidence or information.

(c) Suppose that H is a simple statistical hypothesis, but that $\bar{H}$ is 'composite', i.e. $\bar{H}$ is a logical disjunction $H_1 \vee H_2 \vee H_3 \vee \dots$, of 'simple statistical hypotheses'. (I assume that there are only a countable number of hypotheses, but this is inessential.) Then $P(E|\bar{H})$ can be calculated as

$$P(H_1|\bar{H})P(E|H_1) + P(H_2|\bar{H})P(E|H_2) + \dots,$$

where the Bayesian is prepared to assume subjective values or intervals of values for the probabilities $P(H_1|\bar{H})$, $P(H_2|\bar{H})$, $\dots$. Thus there is a loss of objectivity in $W(H:E)$. We can restore it if we consider the corroboration of H as against (or as compared with, or relative to) H_j for each j separately, that is regard corroboration as a function of j , $W(H:E|H \vee H_j)$, for which an alternative notation is $W(H/H_j:E)$. (Compare Fisher's recommendation, 1956, p. 136, to calculate the likelihood function at each plausible value of the unknown parameter.) This brings out a point that we have not previously emphasized, that the corroboration of a hypothesis is in practice always relative to other hypotheses which might be vague or precise, and this is also true for 'tests of significance'. (Compare the very influential paper by Neyman and Pearson, 1928.) When the alternatives are very general then the test has wider applicability than when they are precise and special, but it is less sensitive: a swings-and-roundabouts effect. If we take as an alternative to H another very similar hypothesis, then we cannot expect to get much corroboration for or against H . If then it is too difficult to estimate the initial probabilities of the hypotheses, it is sensible to interpret corroboration not as a single (scalar) number, but as a function of H_j as we said before.

Similarly, if H is also composite and is a disjunction of hypotheses H_{0i} , whose probabilities we are not prepared to assume we could achieve objectivity by regarding the corroboration as a function of *two* variables, namely $W(H_{0i}/H_j:E)$. (Compare Fisher, 1922, p. 326.) But this explicatum as a function is also sensible for the Bayesian even if some of the *sub-hypotheses* are composite. (It can of course be written as the difference of two functions each of one variable.) The disadvantage of the Bayesian position is its subjectivity, its advantage here is that it enables us to decrease the number of components in the corroboration, even to the extent of reducing it to a scalar. We achieve a greater reduction of the data at the sacrifice of some objectivity, both of which are main aims of statistical methods. Thus the Bayesian controversy is unlikely ever to be resolved.

(d) 'Weight of evidence' is a sharp concept in terms of which to discuss Hempel's paradox of confirmation. This is shown in a few sentences by Good (1966c), but see also Hempel (1967) and Good (1968).

(e) Opponents of the Bayesian position will say that Axiom 1 is acceptable only when the 52670 probabilities all exist, or equivalently the seven probabilities of Theorem 1 exist. When these opponents accept the existence of these probabilities they would I think accept information and weight of evidence as the appropriate explicata. Otherwise I think they will regard the explicatum $W(H_{0i}/H_j:E|G)$ as sensible (as a function of two variables) but only when all the hypotheses H_{0i} and H_j are simple statistical hypotheses. Their thoughts will also turn to the *log-likelihood-ratio*,

$$\log \{ \max_i P(E|H_{0i}, G) / \max_j P(E|H_j, G) \},$$

but, although this is a useful *statistic*, it will not serve as an explicatum for corroboration. (Generalisations of the likelihood-ratio have been proposed; see, for example, Good 1960d, 1967c, but these too cannot serve as strict measures of corroboration.)

(f) On the other hand the Bayesian can say to the non-Bayesian 'Here is a measure of corroboration based on some subjective judgments of probabilities: perhaps you would like to use it as a *statistic* for significance testing if you can compute its distributions analytically or by Monte Carlo methods'. This is an example, and I think a fruitful one, of a Bayes/non-Bayes compromise. (See, for example, Good, 1965, Index.)

(g) Popper (1959, p. 418) says that no formula for $C(H:E)$ can be interpreted as the degree of corroboration of H provided by E unless E reports the results of our sincere efforts to overthrow H , and goes on to say that the requirement of sincerity cannot be formalised. Thus Popper is a subjectivist although he objects to subjective probabilities. (Incidentally he makes use of probabilities of hypotheses in his discussion of corroboration.) Although I consider subjective probabilities to be essential in statistics, science, and life, I think Popper is a little *too* subjective in his insistence on sincerity in this context. Sincerity is useful in the design of experiments, but I can see no objection to regarding $W(H:E)$ as the degree of corroboration of H provided by E , whatever the intentions of the experimenter. The desiderata that led to the explicatum made no reference to sincerity.

Of course we are often concerned with whether the odds of H have gone up or down, allowing for *all* the available evidence. (For a proof that it is rational to use all the evidence we have, and to acquire new evidence when it is free, see Good, 1966d, and the not quite complete anticipations of Raiffa and Schlaifer, 1961, p. 89, and Lindley, 1965, p. 66.) When we are concerned to find new evidence F we should design our experiments to be

informative and this will often involve a sincere effort to overthrow H , for example, by testing it in new kinds of ways.

When we are primarily concerned with testing *one* hypothesis H , we should usually try to design new experiments to ensure that at least one of the mathematical expectations

$$w = \mathcal{E}_F\{W(H:F|E)|H\}$$

and

$$w' = \mathcal{E}_F\{W(\bar{H}:F|E)|\bar{H}\}$$

is large, if the costs are not too high. Now H often has fairly precise implications whereas $\bar{H}$ is liable to be highly composite. Moreover $\bar{H}$ is a disjunction of hypotheses many of which strongly resemble H in their implications for experiments of reasonable cost. The total initial probability of the sub-hypotheses that strongly resemble H is usually not negligible in comparison with that of H . (For example, if H states that all A's are B's, one of the 'close' hypotheses is that between 98 and 99 per cent of A's are B's.) Hence w is not usually very large and w' is usually much larger. This I think is the primary reason why experiments in which overwhelming refutation is 'risked' are more often used than those in which a large positive corroboration can be expected.

But there are plenty of experiments whose main function is the discrimination between a few more or less equally precise hypotheses $H_1, H_2, \dots, H_n$, having initial probabilities that are not very unequal. Consider, for example, the experiments on the deflection of light grazing the sun which were largely for *discriminating* General Relativity from 'Newtonian' physics (here $n = 2$ or 3), or experiments for testing which of a few drugs is the most efficacious. For these experiments it is not necessarily true that we are sincerely trying to refute a particular hypothesis. It might in fact often be sensible to design our experiment so that as many as possible of the expected weights of evidence,

$$w_{ij} = \mathcal{E}\{W(H_i/H_j:E)|H_i\} \quad (i, j = 1, 2, \dots, n; i \neq j)$$

are large. It is not necessarily a matter of sincerely trying to rule out the hypotheses one by one: it might be more economical to try to rule out several of them at once, and perhaps to corroborate just one of them, all in a single well planned experiment, especially in agricultural statistical experiments.

When utilities and costs of rival experiments can be assigned, a rational design maximises 'expected utility'; but when the emphasis is on inference or explanation the utilities can reasonably be interpreted as weights of evidence or as amounts of information. The maximisation of

expected strong-sense explanatory power comes to the same as that of expected information in *experimental design*.

Once an experiment has been performed the maximisation of I or D does not make sense, but that of D' does. Thus the idea of *maximising D' , or its expectation, leads to considerable unification of statistical methods*, and should perhaps be regarded as the real justification for the maximising of expected information in experimental design (when it is justifiable).

5. SIMPLICITY, COMPLEXITY AND THE SHARPENED RAZOR

Very reasonable desiderata can be formulated for 'complexity' without assuming a relationship with probability. The approach used here strongly resembles the treatment of 'information' in Good (1966b).

We assume that our propositions can be stated in finite terms.

Denote the degree of complexity of H given G by $\kappa(H|G)$, where we insist that G should be logically possible. We make the following assumptions:

(a) If H is logically possible, then $\kappa(H|G)$ is a real non-negative number and is not always zero;

(b) $\kappa(H|H) = 0$;

(c) $\kappa(H.K|G) \geq \kappa(H|G)$;

(d) $\kappa(H.K|G) = f\{\kappa(H|G), \kappa(K|H.G)\}$, where f is a fixed continuous function;

(e) If H and K are (or have been proved to be) logically equivalent, then $\kappa(H|G) = \kappa(K|G)$.

From (d) we can deduce that f is 'associative', that is that $f(x, f(y, z)) = f(f(x, y), z)$, and thence by the theorem of the generalised slide rule (due to Abel and Aczél) that $f(x, y)$ is of the form $g^{-1}\{g(x) + g(y)\}$, where g is continuous and increasing. By convention, without affecting inequalities between complexities, we can change the notation $g(\kappa(H|G))$ to $\kappa(H|G)$ and infer that $\kappa(H.K|G) = \kappa(H) + \kappa(K|H.G)$, an additive property which is reasonable in itself. If H and K are in an intuitive sense mutually irrelevant given G we get $\kappa(H.K|G) = \kappa(H|G) + \kappa(K|G)$. Consider the case where K is logically impossible. Since all logically impossible propositions are logically equivalent, we have $\kappa(H.K|G) = \kappa(K|G)$. But, by assumption (a), $\kappa(H|G)$ is not always zero. Therefore the complexity of a logical impossibility is infinite.

It could be argued against this that ' $0 = 1$ ' is a brief and simple statement. But it implies all other statements and therefore, by axiom (c), is at least as complex as any of them. It is the *meaning* of ' $0 = 1$ ' that is infinitely complex, not the sequence of symbols expressing it. This difficulty would

not arise if we were measuring the complexity of physical structures, as in the proposal put into Socrates' mouth by Valéry (1921), but it is a serious objection to the proposal that the complexity of a proposition should be measured by the length of the shortest sentence expressing it in some ideal language. Perhaps the difficulty could be met by having a *syntactic* rule in the ideal language making logical contradictions illegal. Precautions should also be taken to prevent logical antinomies.

The null sentence, the sentence of zero length, says nothing and expresses only a tautology, and is of zero complexity. All (finitely expressible) propositions that are logically possible have finite complexity, by assumption (a). If H and K are mutually irrelevant they can be economically expressed by abutting their individual economical expressions. Thus there is much to be said for Valéry's suggestion.

The desiderata assumed here for $\kappa(H|G)$ coincide with those implied for $I(H|G) = I(H:H|G)$ by Good (1966b). Also the shortest-sentence idea seems equally appropriate for measuring information. Consequently it is natural to identify complexity with amount of (semantic) information even before any reference to probabilities is made. We can reasonably go further and define $\kappa(E:H|G) = I(E:H|G)$ to mean the (amount of) complexity that is *subtracted* from E by H given G . In this manner, complexity, amount of information, and weak-sense explanatory power become completely synonymous.

'Ockham's razor' states that if two hypotheses H and H_1 explain the facts equally, meaning $P(E|H) = P(E|H_1)$, then the simpler of the two is to be preferred. In accordance with our explicatum for complexity, this preference is equivalent to the choice of the more probable hypothesis. The most probable hypothesis is a tautology but this will not usually be a candidate since Ockham's razor is intended to apply only when $P(E|H) = P(E|H_1)$, i.e. when $D(H:E) = D(H_1:E)$. We need a sharper razor for the choice between hypotheses when this condition is not satisfied. In view of our earlier discussions we call the following principle the 'Sharpened Razor': *Choose the hypothesis whose strong explanatory power is the greater, or the greatest.* We have already considered an application of the Sharpened Razor to binomial sampling in Section 3, and it gave very sensible results. If we were to use the unsharpened razor when $D(H:E) \neq D(H_1:E)$ it would be as if we had taken the constant γ equal to unity, which, to put it bluntly, is unreasonable, as we saw in Section 3.

'Simplicity' should clearly be taken as a strictly decreasing function of complexity. If H and K are mutually irrelevant and equally simple it is natural to say that H is twice as simple as $H.K$. If we accept this, with its obvious extension to more than two propositions, then we are committed to

defining simplicity as the reciprocal of complexity. But complexity is slightly the simpler notion owing to its additivity.

Rothstein (1958, p. 45) defines the complexity of a theory with respect to its function (purpose) as the excess of the entropy of the theory over that of its function. It is not clear what these entropies are since the relevant ensembles are undefined, but his suggestion is a less formal and partial anticipation of our $\kappa(E:H|G)$. A vague interpretation of the entropy of the observations would be the expected amount of information per dollar that should rationally be spent for the *optimal* strategy of experimentation.

Jeffreys (1957, p. 131) mentions that the Copernican epicycle theory H , with 33 terms, would explain all the motions E of the planets to the first order in the eccentricities. No doubt with enough terms it would explain the motions even better than Kepler's laws K . But the Sharpened Razor would still tell us to prefer Kepler's laws provided they were expressed in the form that the orbits are approximately ellipses, the degree of approximation being suitably formulated. This is intuitively clear in spite of the fact that the formulation of the degree of approximation has (a) not been done, (b) would be *ad hoc*. For the additional complexity thereby introduced in K would be much smaller than that introduced in H (by numerous new parameters), whereas $I(K:E)$ would not be much smaller than $I(H:H)$. This argument succeeds if γ is not very small. Examples like this should enable us eventually to circumscribe the value of γ more precisely, and perhaps $\gamma = \frac{1}{2}$ will be found to be acceptable.

It is of interest to ask why Popper argued that a simple theory is one of low logical probability. Partly perhaps because he had noticed examples of the kind brought to my attention by Agassi, and discussed in Section 3; and largely because of his contention that the logical probability of a universal law is usually zero in an infinite universe. His arguments for this seem to me to be fallacious: see, for example, Good (1960c) and Popper's attempt (Popper, 1959, p. 371) to refute a 'Bayesian' justification of a kind of induction (Jeffreys, 1961, §1.62; compare Keynes, 1921, chap. 20). On a careful mathematical examination of this attempted refutation it can be seen to prove only that *infinitely complicated* propositions have probability zero, which we do not dispute.

A modified version of Popper's strictures on induction seems to me to be valid however since we are not often concerned with probabilities of single hypotheses (or theories). It is the *relative* probabilities of pairs of hypotheses which enter into the formula

$$\log O(H/H_1|E) = \log O(H/H_1) + W(H/H_1:E),$$

and which are more relevant to scientific inference. The log-odds can be written as the difference between two complexities if preferred.

If complexity could be measured without an appeal to probability we could define probability as $e^{-\kappa/\lambda}$, where λ depends only on the unit of measurement of complexity. This is of the form of the density of Boltzmann's distribution for a physical system if complexity is identified with physical entropy.

Appendix A. Let f be a real function of $P(E|H)$ and $P(E|\bar{H})$ that, together with $P(H)$, mathematically determines $P(H|E)$. (The functional symbols f and g in this appendix are not intended to be related to previous usages, nor to those in Appendix B.) We stated at the beginning of Section 3 that this function must be a monotonic function of $W(H:E)$ and a proof of this assertion will now be outlined.

By assumption, there is a function g such that

$$P(H|E) = g[f(P(E|H), P(E|\bar{H})), P(H)]$$

is identically true. Writing $P(H) = x$, $P(E) = y$ and $P(H|E) = z$, we have

$$z = g\left[f\left(\frac{zy}{x}, \frac{y(1-z)}{1-x}\right), x\right].$$

Thus g is mathematically independent of x , and so must depend on the value taken by f . Hence f must be mathematically independent of y , so it must depend only on the ratio of its arguments. This function must be monotonic otherwise it would take the same value for two different values of z , and the above equation would be impossible. Thus f is a monotonic function of $O(H|E)/O(H)$ or equivalently of $W(H:E)$.

Appendix B. $D'(H:E)$ was assumed to be a function of $P(H)$ and $D(H:E)$, equivalently of $P(H)$ and $P(H|E)$. Denote this function by $D'(H:E) = (P(H), P(H|E)) = f(x, y)$ say. We assumed further that $D'(H.K:E.F)$ was some function of $D'(H:E)$ and $D'(K:F)$ when H and E have nothing to do with K and F . Thus there is a function g such that

$$f(xx', yy') = g(f(x, y), f(x', y')).$$

Let $xx' = X$, $yy' = Y$, so that

$$f(X, Y) = g(f(x, y), f(X/x, Y/y)).$$

I shall now make use of the differentiability assumption of Axiom 1 since this leads to a quick solution. (Continuity alone is probably sufficient, in accordance with the usual behaviour of functional equations.) Denote partial derivatives with respect to first and second arguments by suffixes 1 and 2. Then we have

$$0 = f_1(x, y)g_1 - Xx^{-2}f_1(X/x, Y/y)g_2$$

$$0 = f_2(x, y)g_1 - Yy^{-2}f_2(X/x, Y/y)g_2$$

Therefore

$$\frac{xf_1(x,y)}{yf_2(x,y)} = \frac{x'f_1(x',y')}{y'f_2(x',y')}$$

so that the left side must be independent of x and y . Call it A . Then

$$\frac{\partial f}{\partial \log x} = A \frac{\partial f}{\partial \log y}$$

Therefore $f(x,y)$ is a function of $A \log x + \log y$. The form given for $D'(H:E)$ in the body of the paper now follows from the other assumptions made there.

REFERENCES

- Bar-Hillel, Y. and Carnap, R. (1953), 'Semantic Information', this *Journal*, **4**, 147-157.
 Carnap, R. (1950), *Logical Foundations of Probability*, Chicago: Routledge and Kegan Paul.
 Fisher, R. A. (1922), 'On the mathematical foundations of theoretical statistics', *Phil. Trans. Roy. Soc. Ser. A.*, **222**, 309-68.
 Fisher, R. A. (1956), *Statistical Methods and Scientific Inference*, New York: Hafner.
 Good, I. J. (1950), *Probability and the Weighing of Evidence*, London: Griffin; New York: Hafner.
 Good, I. J. (1956), 'Some notation and terminology in information theory', *Proc. Institution Elec. Engrs.*, Part C, **103**, 200-4 = Monograph 155 R.
 Good, I. J. (1959a), 'Kinds of probability', *Science*, **129**, 443-7.
 Good, I. J. (1959b), 'Lattice structure of space-time', this *Journal*, **9**, 317-19.
 Good, I. J. (1960a), 'Gravitational frequency shift', letter in *New Scientist*, **7** (Feb. 25), 483.
 Good, I. J. (1960b), 'Weight of evidence, corroboration, explanatory power, information and the utility of experiments', *J. Roy. Statist. Soc. B*, **22**, 319-31. Corrigenda forthcoming (but the theorems remain valid).
 Good, I. J. (1960c), review of Popper (1959); *Math. Rev.* **21**, 1171-3.
 Good, I. J. (1960d), Contribution to the discussion of a paper by E. M. L. Beale, *J. Roy. Statist. Soc.*, ser. B, **22**, 79-82, esp. p. 80.
 Good, I. J. (1960/61), 'The paradox of confirmation', this *Journal*, **11**, 145-8; **12**, 63-64.
 Good, I. J. (1961), 'A causal calculus', this *Journal*, **11**, 305-18; **12**, 43-51; **13**, 88.
 Good, I. J. (1962), 'How rational should a manager be?' *Management Science*, **8**, 383-93; reprinted with improvements in *Executive Readings in Management Science* (ed. M. K. Starr); New York: Macmillan, 1965, 88-98.
 Good, I. J. (1964), 'Measurement of decisions', in *New Perspectives in Organization Research* (eds. W. W. Cooper, H. J. Leavitt, and M. W. Shelley, II); New York, Wiley, 391-404.
 Good, I. J. (1965), *The Estimation of Probabilities: an essay on modern Bayesian methods*, M.I.T. Press.
 Good, I. J. (1966a), 'How to estimate probabilities', *J. Inst. Math. Applics.* **2**, 364-83.
 Good, I. J. (1966b), 'A derivation of the probabilistic explication of information', *J. Roy. Statist. Soc. Ser. B*, **28**, 578-81.
 Good, I. J. (1966c), 'The white shoe is a red herring', this *Journal*, **17**, 322.
 Good, I. J. (1966d), 'On the principle of total evidence', this *Journal*, **17**, 319-21.
 Good, I. J. (1967a), Contribution to the discussion of S. F. Buck and A. J. Wicken, 'Models for use in investigating the risk of mortality from lung cancer and bronchitis', *Applied Statistics* (forthcoming).
 Good, I. J. (1967b), 'A five-year plan for automatic chess', in *Machine Intelligence, II* (ed. E. Dale and D. Michie; Edinburgh: Oliver and Boyd), 89-118.

- Good, I. J. (1967c), 'A Bayesian significance test for multinomial distributions', *J. Roy. Statist. Soc. Ser. B* (forthcoming).
- Good, I. J. (1967d), 'Probability or support?', a review of Hacking (1965), *Nature*, **213**, 233-4.
- Good, I. J. (1968), 'The white shoe, *qua* herring, is pink', this *Journal*, to be published.
- Good, I. J. and Toulmin, G. H. (1967), 'Coding theorems and weight of evidence', *J. Inst. Math. Applics.* (forthcoming).
- Hacking, I. (1965), *Logic of Statistical Inference*, Cambridge: University Press.
- Hempel, C. G. (1967), 'The white shoe: no red herring', this *Journal*, **18**, 39-40.
- Jeffreys, H. (1936), 'Further significance tests', *Proc. Cambridge Philos. Soc.* **32**, 416-45.
- Jeffreys, H. (1957), *Scientific Inference*, 2nd edn.; Cambridge: University Press.
- Jeffreys, H. (1961), *Theory of Probability*, 3rd edn.; Oxford: University Press.
- Kapp, R. O. (c. 1961), A lecture to the Brit. Soc. Philos. Science, London.
- Keynes, J. M. (1921), *A Treatise on Probability*, London: Macmillan.
- Kullback, S. (1959), *Information Theory in Statistics*, New York: Wiley.
- Lindley, D. V. (1956), 'On a measure of the information provided by an experiment', *Ann. Math. Statist.* **27**, 986-1005.
- Lindley, D. V. (1965), *Introduction to Probability and Statistics*, Pt. 2, Cambridge: University Press.
- McVittie, G. C. (1965), *General Relativity and Cosmology*, New York: Wiley.
- Neyman, J. and Pearson, E. S. (1928), 'On the use and interpretation of certain test criteria for the purposes of statistical inference', *Biometrika*, **20A**, 175-240 and 263-94.
- Poincaré, H. (1913/1963), *Dernières Pensées*, Paris: Flammarion; Eng. tr., New York, Dover, 1963.
- Popper, K. R. (1959), *The Logic of Scientific Discovery*, London: Hutchinson; New York: Basic Books.
- Raiffa, H. and Schlaifer, R. (1961), *Applied Statistical Decision Theory*, Boston: Harvard University Press.
- Rothstein, J. (1958), *Communication, Organization, and Science*, Indian Hills, Colorado: Falcon's Wing Press.
- Shannon, E. C. and Weaver, W. (1949), *A Mathematical Theory of Communication*, Urbana, Illinois.
- Turing, A. M. (1940), private communication.
- Valéry, P. A. (1921), 'Eupalinos ou l'Architecte-Dialogue des morts', *La Nouvelle Revue Française*, no. 90, 237-85. Reprinted in *Oeuvres de Paul Valéry*, II (Editions Guillimard, 1960), 79-147. (On p. 109 a figure is called 'geometric' if it can be traced by motions which can be expressed in few words.)
- Zeeman, E. C. (1965), 'Topology of The Brain', in *Mathematics and Computer Science in Biology and Medicine*, London: H.M.S.O., 277-92.

Virginia Polytechnic Institute

Discussions

THE STRAIGHT AND NARROW RULE OF INDUCTION

A Reply to Dr Bub and Mr Radner

I

In [5] I tried to show that a probabilistic form of the 'straight rule of induction' (along with a wide class of related rules) is inconsistent. It seems, therefore, that anyone who wishes to defend this form of the straight rule (or any of the other rules) is obliged to show either that (a) the derivation I gave is *invalid* (this is what Professor Mackie tried to do in [4]); or that (b) the derivation is valid, but the rule can be stated in such a way that my argument is inapplicable to it.

In [1] Bub and Radner nowhere state clearly whether it is (a) or (b) that they wish to maintain. Indeed, in their penultimate paragraph—which, with its footnote, is the only part of their comment that is not completely trivial—they seem to start by arguing for (a), but quite clearly finish by arguing for (b). I shall try to show, however, that they have nothing whatever to say in favour of (a); and I shall thereafter (in II) assume that (b) is the thesis of their note.

The passage, referred to above, that is *prima facie* in favour of (a), says (the italics are mine):

Miller's account is persuasive because his notation leads us quickly and *deceptively* through the steps of his proof. For his equation (1)

$$P\left(A, E_r^a\right) = r$$

which expresses the Straight Rule, there is a tendency to read the subscript of ' E_r^a ', to obtain the value of the function P . This habit *seems* innocuous. However, . . .

Now, to my mind, not only is the 'habit' they describe innocuous; it is actually essential if we wish to do mathematics (it is the 'habit' of not substituting *different* constants for different occurrences of the *same* free variable). Bub and Radner, on the contrary, seem to hint (without being too explicit about it) that my derivation is *invalid* because I unconsciously fall prey to this 'habit', as a consequence of which 'the function P becomes multi-valued and allows . . . [me] to derive a contradiction'. But this is, as I have already explained, a misunderstanding, and, incidentally, at odds with the great bulk of their note—the elaboration of what they call 'the central thesis of Mackie's note'.¹ In view of this I conclude that Bub and Radner do not really wish to argue for (a).

¹ See p. 63 of [1]. Bub and Radner go on to say of this supposed thesis of Professor Mackie's: 'Mackie pointed out, quite correctly, that this paradox has nothing to do with the concept of probability. In fact it arises quite trivially whenever a function is not well-defined.' But, *in this sense*, Russell's paradox, to take a simple example, 'has nothing to do with' set theory. In fact, 'it arises quite trivially' whenever a theory contains (say) a

II

Before discussing Bub and Radner's arguments in favour of (b), I wish to say a little about the general problem that faces those defenders of the straight rule who admit the validity of my derivation.

Briefly stated, it is this. If we add to my (1) the restriction that only *numerals*¹ may be substituted for '*r*', then we obtain a version of the straight rule to which my argument appears to be inapplicable, and which we may well *conjecture* to be consistent. We can even develop it into the following recursive version of (1), a rule that, in view of its extreme narrowness of application, I shall call the *narrow rule of induction*.

$$(NR) \quad P(A, E_\lambda^a) = \lambda;$$

if E_u^a is logically equivalent to E_λ^a , then $P(A, E_u^a) = \lambda$.

Unfortunately, this rule is so patently *ad hoc* that we cannot be expected to take it seriously.²

relation symbol ' ϵ ', a theorem ' $(x)(x\epsilon a \leftrightarrow x\epsilon x)$ ', and a rule allowing the instantiation of ' x ' by ' a '. Russell's paradox, however, was important (and not 'trivial'), because it showed that Frege's *Grundgesetze der Arithmetik* was such a theory (I ignore the fact that Frege's terminology bore little relation to that given above). Similarly, my note was not intended as a reiteration of the utterly obvious fact that a theory which allows the *identification of different* values of a multi-valued function is inconsistent. Rather, what I tried to show (although I did not state it in precisely this way) was the perhaps somewhat surprising fact that any theory that contains my (1) *is such a theory*.

Thus Bub and Radner's charge (see the paragraph containing their equations (15a) and (15b)) that '*relative to the derivation of the paradox*, Miller's notation is . . . redundant (with respect to the proposition *A*)' is not exactly a criticism. I agree with the remark of course. But this does not mean that I would have done well to omit '*A*'. For it is of no interest that statements like their (15b) (which anyway violates the addition law) happen to be inconsistent. What is perhaps interesting is that statements like my (1) (their (15a)) *have been seriously put forward and thought to be consistent*. To give only two clear-cut examples: on p. 39 of [7], K. R. Popper proposes a formula (PP) which he claims is *logically true*, but which is clearly inconsistent (I am indebted to Professor Popper himself for drawing my attention to this statement); and rule (3g) on p. 389 of C. G. Hempel's [3] is virtually a translation of (1) into his terminology.

¹ Or, more generally, what Carnap calls 'L-determinate [number] designators'; cf. [2], pp. 69-73.

For simplicity I shall use Greek letters as syntactic (dummy) variables for numerals. A formal expression containing the symbol ' λ ' is to be read with every occurrence of ' λ ' replaced by λ ; thus if λ is the numeral '0', the name ' E_λ^a ' is the name ' E_0^a '. The statement ' $x = \lambda$ ' will always mean that λ is a name for the number x , and never that x is the same numeral as λ ; similarly, ' $\lambda = \mu$ ' will mean that λ and μ are numerals for the same number. I shall moreover take it for granted that λ may be substituted for the number variable '*r*' in ' E_r^a ' only if $0 < \lambda < 1$ (for otherwise all versions of the straight rule become hopelessly inconsistent). This is most economically done by demanding that our information statements always be consistent.

The letter '*u*' will also be used as a syntactic variable for number designators in general.

² I am grateful to Professor Popper for pointing out to me not only that the narrow rule seems to avoid the inconsistency, but also that, because of its reference to numerals, it is

The problem, therefore, is to find a (consistent) statement of the straight rule which not only avoids all reference to numerals, but which also is itself not too *ad hoc*, and, if possible, is of wider application than the narrow rule. In particular, as Bub and Radner emphasise, the restatement of the rule, to be consistent, must preserve the univalence of the function P .

My thesis is that Bub and Radner have not found such a restatement of the straight rule.

For the purpose of this reply, I intend to direct my attention exclusively at the rule advanced by Bub and Radner in their footnote on p. 66, since this rule,

$$(BR) \quad P\left(A, E_{f(p(a))}^a\right) = f\left(P\left(A, E_{f(p(a))}^a\right)\right),$$

is intended by them to be a generalisation of the rule stated in their penultimate paragraph: '[interpret] Miller's (1) as stating that " E_r^a " designates an equation which, if yielding a value of $p(a)$ by solution, assigns that value to P '; and also of the rule expressed by (20).¹

Before I start to criticise (BR), I should like to say that it seems to me a most interesting and ingenious attempt to save the straight rule from the paradox of information, and that I am very grateful to Bub and Radner for suggesting it. Indeed, I am quite sorry that the rule leads almost immediately to related contradictions.

My argument consists of four steps:

- (i) (BR) is inconsistent;
- (ii) This inconsistency can be avoided only by an undeniably *ad hoc* stratagem, leading to a rule I will call (BR');
- (iii) (BR') is inconsistent in the usual (unrelativised) systems of probability theory;
- (iv) This inconsistency can be avoided only by so changing (BR') that it becomes equivalent to the narrow rule.

(i) Let us introduce the following *ad hoc* terminology. We shall call a (number) designator ' $p(a)$ -determinate' if, roughly speaking, it can be defined solely in terms of ' $p(a)$ '. More precisely, u is $p(a)$ -determinate if the designator obtained from u by replacing ' $p(a)$ ' throughout by a numeral is 'L-determinate' (so an L-determinate designator is $p(a)$ -determinate). Thus, for example, ' $[p(a)]^2$ ' is $p(a)$ -determinate. 'Avogadro's number' and 'the number of angels that can dance on the head of a pin' are $p(a)$ -indeterminate.²

unacceptably *ad hoc*, and, apparently, has no precedent in mathematics. I have no doubt, however, that some defenders of the straight rule will find that it strings along nicely with their 'inductive intuition'.

¹ Whether these rules are as closely connected as Bub and Radner intend is a point that I cannot discuss here. But it is worth noting that (20), *as it stands*, is obviously inconsistent, since we have only to substitute 'o' for 'r' to get back to my (1).

² Any 'mathematical' designator will here be treated as L-determinate (although it would not affect my argument if I restricted L-determinateness even to computability). This can be done by expanding the class of logical signs to include all mathematical symbols;

My claim is that if we substitute for ' $f(p(a))$ ' in (BR) a $p(a)$ -indeterminate designator (so that f is a constant function), the rule quickly leads to a contradiction.

This is easily seen as follows. (For typographical reasons we will use ' e ' as an abbreviation for such a designator—for example, the designator 'the ratio of the mass of the electron to the mass of the proton'; so ' e ' is $p(a)$ -indeterminate.) Substituting ' $e-1$ ' for ' r ', we have, by (BR),

$$P\left(A, E_{e-1}^a\right) = e-1,$$

since $e-1$ is the same function of $P\left(A, E_{e-1}^a\right)$ that it is of $p(a)$. But all logical probabilities lie between 0 and 1; so we obtain

$$0 \leq e-1 \leq 1;$$

that is,

$$1 \leq e \leq 2,$$

which, being a factual statement, shows that the rule (BR) is unacceptable. But we can go a bit further by similarly substituting ' $e+1$ ' for ' r ' and arriving at

$$-1 \leq e \leq 0.$$

Since these two sets of bounds for e are disjoint, we have a contradiction.

(ii) Thus the free substitutability of $p(a)$ -indeterminate designators ' e ' for ' $f(p(a))$ ' makes (BR) inconsistent. And the least drastic way out of this, I think, is to require that ' e ' may only be substituted for ' $f(p(a))$ ' if $0 \leq e \leq 1$. But the resulting rule cannot be used for calculating anything like logical probabilities. For suppose the statement

$$P\left(A, E_e^a\right) = e$$

is obtained by means of the rule. Since it has the consequence

$$0 \leq e \leq 1,$$

which is *factual* (except in certain special cases; for example, when ' e ' abbreviates ' $p(b)$ '), it *must itself be factual* (and so neither 'analytic' nor conventional).

To put the point even more strongly: to be able to write down (in principle) all the logical probability statements derivable from (BR), we would have to know the value of every $p(a)$ -indeterminate designator to an arbitrarily high degree of precision. For, whatever numeral λ may be (apart from numerals for 0), we would

have to know whether $0 \leq e/\lambda \leq 1$ in order to calculate $P\left(A, E_{e/\lambda}^a\right)$; that is, whether

(see [2], pp. 85–88; for the absence of a sharp boundary between logical and non-logical signs, see [8], pp. 418 f.). Furthermore (and this again is not essential), I shall include

' p ' amongst the logical signs, so that, for example, $E_{1-p(-a)}^a$ (that is, the statement

' $p(a) = 1 - p(-a)$ ') is logically true, whilst $E_{p(b)+2}^a$ is logically false.

The only really important point is that 'logic' should mean the same here as it does in the statement of the narrow rule.

$0 \leq e < \lambda$. (This is true even of designators like ' $p(b)$ '.) But a straight rule so dependent on matters of fact could hardly be of much use, for example, in what is called 'inductive logic'.

Thus, to retain any significance (BR) must be so restricted that only $p(a)$ -determinate designators may be substituted for ' $f(p(a))$ '. Let us call the rule amended in this way (BR').

(iii) I think that (i) and (ii) together are sufficient to show that (BR') is not much improvement on the narrow rule. For the restriction mentioned in the previous paragraph is as *ad hoc* as the restriction to numerals in the narrow rule.

However, it is of some interest to see in what respects the narrow rule and

(BR') differ. If our evidence statement is E_u^a , where u is $p(a)$ -indeterminate, then our two rules will agree (in being inapplicable, if they are consistent); and if the evidence is logically equivalent to E_λ^a where λ is a numeral, then also the rules

agree. For example, since $E_{[p(a)]^2 + \frac{1}{2}}^a$ is logically equivalent to $E_{\frac{1}{2}}^a$, we have by the narrow rule

$$P\left(A, E_{[p(a)]^2 + \frac{1}{2}}^a\right) = \frac{1}{2};$$

and by (BR') we obtain

$$P\left(A, E_{[p(a)]^2 + \frac{1}{2}}^a\right) = \left(P\left(A, E_{[p(a)]^2}^a\right) + \frac{1}{2}\right)^2 + \frac{1}{2},$$

which gives the same result.

But consider the equation $E_{[p(a)]^2}^a$. It is easy to see that this is equivalent to $E_0^a \vee E_1^a$ (I use the connectives in their usual senses), so that the narrow rule gives us no value for $P\left(A, E_{[p(a)]^2}^a\right)$. Rule (BR'), however, can be used, and it leads to some rather surprising results.

For, first of all, we get immediately from (BR')

$$P\left(A, E_{[p(a)]^2}^a\right) = \left(P\left(A, E_{[p(a)]^2}^a\right)\right)^2$$

which is equivalent to the disjunction

$$(\alpha) \quad P\left(A, E_{[p(a)]^2}^a\right) = 0 \quad \text{or} \quad P\left(A, E_{[p(a)]^2}^a\right) = 1.$$

Now (α) itself is fairly absurd. It tells us, for example, that the logical probability that a coin will fall heads on the next throw, given that it always falls the same way up, is either 0 or 1. And if logical probability is interpreted as measuring something like the support one statement gives to another, (α) (and, even more so, (α') below) can easily be developed into a form of the paradox of ideal evidence (cf. [6], pp. 406-409).

This is, I think, a fairly severe criticism of (BR'). But it is overshadowed by the fact that we can use (α) to derive a contradiction.

For reasons of simplicity, I shall carry out the derivation in the usual (Kolmogorov) system of probability theory, where $P(B, C)$ is defined by

$$P(B, C) = P(B \wedge C)/P(C),$$

so that it is meaningless when $P(C) = 0$.¹

Thus, given that 'o', '1', ' $[p(a)]^2$ ' can all be substituted for ' $f(p(a))$ ' in (BR'), we at once find that

$$(*) \quad P(E_o^a) \neq 0; \quad P(E_I^a) \neq 0; \quad P(E_{[p(a)]^2}^a) \neq 0.$$

From (*) and (BR') it follows that $P(A) \neq 0$. For, otherwise,

$$P(A, E_I^a) = P(A \wedge E_I^a)/P(E_I^a) < P(A)/P(E_I^a) = 0,$$

which contradicts (BR'). Thus $P(B, A)$ is defined for all B .

Now the following is a theorem of the probability calculus which is easily proved under these conditions from Bayes's theorem and the addition law.

If B and C are incompatible, then

$$P(A, B \vee C) = P(A, B)P(B, B \vee C) + P(A, C)P(C, B \vee C).$$

Since $E_{[p(a)]^2}^a$ is equivalent to the disjunction $E_o^a \vee E_I^a$, and E_o^a and E_I^a are incompatible, we have at once that

$$P(A, E_{[p(a)]^2}^a) = P(A, E_o^a)P(E_o^a, E_o^a \vee E_I^a) + P(A, E_I^a)P(E_I^a, E_o^a \vee E_I^a).$$

But from (BR') we have

$$P(A, E_o^a) = 0; \quad P(A, E_I^a) = 1,$$

and so

$$(\alpha') \quad P(A, E_{[p(a)]^2}^a) = P(E_I^a, E_o^a \vee E_I^a);$$

and since, according to (α), $P(A, E_{[p(a)]^2}^a)$ is either zero or one, we can conclude

that $P(E_I^a, E_o^a \vee E_I^a)$ is zero or one.

Using the addition theorem in the form

$$P(E_o^a, E_o^a \vee E_I^a) = 1 - P(E_I^a, E_o^a \vee E_I^a)$$

we see that if $P(E_I^a, E_o^a \vee E_I^a)$ is not zero, then $P(E_o^a, E_o^a \vee E_I^a)$ is zero, and vice versa; thus

$$(\alpha'') \quad P(E_I^a, E_o^a \vee E_I^a) = 0 \quad \text{or} \quad P(E_o^a, E_o^a \vee E_I^a) = 0.$$

¹ In the relativised theory of probability given in detail in [6], pp. 326-358, a similar contradiction can be avoided only if we make a ridiculous assignment of absolute probabilities. But the proof of this cannot be given here.

$$\begin{aligned} \text{But } P(E_I^a, E_O^a \vee E_I^a) &= P(E_I^a \wedge (E_O^a \vee E_I^a) / P(E_O^a \vee E_I^a)) \\ &= P(E_I^a) / P(E_O^a \vee E_I^a). \text{ Likewise, } P(E_O^a, E_O^a \vee E_I^a) = P(E_O^a) / P(E_O^a \vee E_I^a). \end{aligned}$$

Since, by (α^*), one of these two is zero, we get the disjunction

$$P(E_O^a) = 0 \quad \text{or} \quad P(E_I^a) = 0$$

which contradicts (*). So (BR') is inconsistent.

(iv) In a similar way, if μ and ν are numerals and $\mu \neq \nu$, we can show that

$$(\beta^*) \quad P(E_\mu^a, E_\mu^a \vee E_\nu^a) = 0 \quad \text{or} \quad P(E_\nu^a, E_\mu^a \vee E_\nu^a) = 0,$$

which leads to the same contradiction. To prove (β^*) we substitute

$$([p(a)]^2 + \mu\nu) / (\mu + \nu)$$

for ' $f(p(a))$ ' in (BR'), and, exactly as above, obtain

$$(\beta') \quad P(A, E_\mu^a \vee E_\nu^a) = \mu P(E_\mu^a, E_\mu^a \vee E_\nu^a) + \nu P(E_\nu^a, E_\mu^a \vee E_\nu^a).$$

The left-hand side is either μ or ν (by (BR')), and we then proceed by cases to obtain (β^*). Thus the contradiction cannot be sidestepped by a simple *ad hoc* restriction on ' $[p(a)]^2$ ' as a possible substituent for ' $f(p(a))$ '.

And, indeed, the only plausible way of avoiding the contradiction is to forbid the substitution for ' $f(p(a))$ ' of any designator u which makes E_u^a a non-linear equation in ' $p(a)$ '. For similar contradictions can easily be derived for equations of any finite degree (and even for equations of infinite degree). But this restriction makes (BR') equivalent to the narrow rule, and it cannot be said that the formulation is any less *ad hoc*.

My conclusion is that Bub and Radner have failed to provide a rule that solves the paradox of information any better than the narrow rule does; and that the paradox cannot therefore be said to be solved.

The possibility is still open, however, of consistently formulating the straight rule in a completely different way. But this would be a topic for a further note.

This reply was written while the author was studying as a University Fellow at Stanford University, with the assistance of a Fulbright travel grant from the U.S.-U.K. Educational Commission. Both forms of support are gratefully acknowledged.

DAVID MILLER
Stanford University

REFERENCES

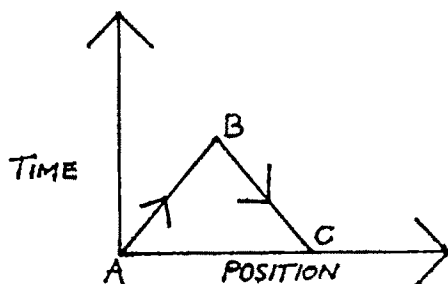
- [1] Bub, Jeffrey, and Radner, Michael, 'Miller's Paradox of Information', this *Journal*, 19, May, 1968, 63-67.
- [2] Carnap, Rudolf, *Meaning and Necessity*, 1947 and later editions, The University of Chicago Press.

- [3] Hempel, Carl G., *Aspects of Scientific Explanation*, 1965, The Free Press, New York.
- [4] Mackie, J. L., 'Miller's So-called Paradox of Information', this *Journal* 17 (August 1966), 144-147.
- [5] Miller, David, 'A Paradox of Information', this *Journal* 17 (May 1966), 59-61.
- [6] Popper, Karl R., *The Logic of Scientific Discovery*, 1959, Hutchinson & Co., London.
- [7] Popper, Karl R., 'The Propensity Interpretation of Probability', this *Journal* 10 (May 1959), 25-42.
- [8] Tarski, Alfred, *Logic, Semantics, Metamathematics*, translated by J. H. Woodger, 1956, Oxford University Press.

THE CONCEPTUAL POSSIBILITY OF TIME TRAVEL

In this paper I will attempt to refute the thesis that it is 'conceptually impossible' for a physical object to 'travel backwards in time'. I take the last-quoted phrase to embody the following ideas, made *familiar* by modern physics:

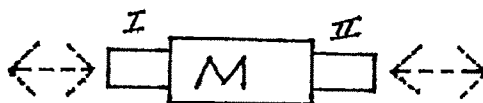
- (1) The trajectory of a physical object is given by its world-line in Minkowski space.
- (2) If an object 'travelled backwards in time', its trajectory would consist of at least two (time-like) segments AB, BC (see below), such that:
 - (a) the object would first go from A to B, for which $\Delta t > 0$.
 - (b) the object 'then' goes from B to C, for which $\Delta t < 0$. ('Then' refers to some 4-dimensional parametrisation of the entire world-line ABC).



My strategy will be to exhibit a possible extension of current physical theory, and then to outline the conditions under which we would *have* to say that some object 'travelled backwards in time', or as I like to say, 'was re-entrant'. Another result will be that the admission of re-entrance does not require some conceptual change in the everyday notion of time, such changes being construed as changes of meaning.

I. The Feynman Algorithm

Let us consider the following piece of quantum-mechanical machinery M:



Following Feynman,¹ we say that a particle at I is in the *entry state*, and that a particle at II is in the *exit state*. M itself forms part of a (time-like) world-tube for which the time co-ordinate for I, t_1 , is less than that of II, t_2 .² We write then, $I(t_1)$, $II(t_1)$, $t_2 - t_1 > 0$. For processes going from $I(t_1)$ to $II(t_2)$ and from $II(t_2)$ to $I(t_1)$, we have $\Delta t > 0$, $\Delta t < 0$ respectively. We write these as $I(t_1) \rightarrow II(t_2)$, $II(t_2) \rightarrow I(t_1)$. Now let us stipulate that at least two kinds of particles exist, electrons (e^-) and positrons (e^+). They differ in charge sign. A typical statement about an electron undergoing some process can be written as $Pe^-(I(t_1) \rightarrow II(t_2))$. Let arrows (possibly double-headed) between such inscriptions be read as 'can be translated into' or 'rewrite'. Then for reasons of energy conservation, Feynman lays down the following translation algorithm:

- (R₁) $Pe^+ \leftrightarrow Qe^-$. In particular, for our machine,
- (R₂) $Pe^+(I(t_1) \leftrightarrow II(t_2)) \leftrightarrow Qe^-(II(t_2) \leftrightarrow I(t_1))$
- (R_{2a}) the sense of every arrow in (R₂) must agree in any given application.

The following translation is forbidden:

$$(FR_1) Te^-(II(t_2) \rightarrow I(t_1)) \leftrightarrow Ue^-(I(t_1) \rightarrow II(t_2))$$

The double arrows in R₁, R₂, show that the algorithm represents a form of *full conventionalism* in the sense that *any* statement about the temporal behaviour of a positron (electron) can be translated into one about that of an electron (positron).³ The algorithm is not at present embodied in a theory, which has essential recourse to re-entrant ($\Delta t < 0$) trajectories, and is of greater predictive power than one using both e^+ and e^- particles and only non-re-entrant ($\Delta t > 0$) trajectories. This is assured by the above rules. In a sense made familiar by Bohr,⁴ it can be argued that we must already know the two particle non-re-entrant formulation before we can translate a given process-description into Feynman's form. The algorithm is then viewed as a convenient fiction, and we never *have* to say that electrons travel backwards in time.⁵ In this form then, Feynman's theory seems philosophically uninteresting.

II. Extension of the Algorithm

Let us suppose that a new microphysical theory was developed, which was far more successful than present-day quantum-electro-dynamics. Assume further that the theory contained the following algorithm for the translation of positron-statements:

- (R'₁) $Pe^+ \rightarrow Qe^-$
- (R'₂) $Pe^+(I(t_1) \leftrightarrow II(t_2)) \rightarrow Qe^-(II(t_2) \leftrightarrow I(t_1))$
- (R'_{2a}) The arrows inside the above parentheses must agree in sense.
- (FR'₁) $Te^-(II(t_2) \rightarrow I(t_1)) \leftrightarrow Ue^-(I(t_1) \rightarrow II(t_2))$.

¹ R. P. Feynman, *Theory of Fundamental Processes* (New York, Benjamin, 1962), ch. 5.

² In special relativity such an assignment of time-coordinates is unique in the sense that if one observer orders world-points P_1, P_2 such that $t_1 < t_2$, then all observers must do so. See J. L. Synge, *Relativity: the Special Theory* (New York, Interscience, 1956), pp. 73-74.

³ I am not sure that this holds for all trajectories, but let us take it as so holding.

⁴ Niels Bohr, *Atomic Theory and the Description of Nature* (Cambridge, Cambridge University Press, 1961).

⁵ Mario Bunge, *Metascientific Queries* (Springfield, Ill., Charles Thomas, 1959), pp. 249-68. Also Henry Margenau, 'Can Time Flow Backwards', *Philosophy of Science*, vol. xxi (1954).

Such a theory would not be fully conventionalistic in the above sense, since while positron-statements can be translated into electron-statements, the inverse translation is ruled out by the arrows. If we interpret R'_1 as saying that all positron-statements *must* be translated into statements about (possibly re-entrant) electrons, then R'_1 can be viewed as a framework-principle stating that there are no positrons. If this statement is generalized to include all particle-antiparticle pairs, and if this generalisation is assumed to hold for the description of all microprocesses, then this principle can be extended to the following 'ontological' claim:

(P) There are no antiparticles.

In this new theory (P) requires us to *replace* all antiparticle statements which at present describe sub-atomic processes, e.g. pair creation, by statements about re-entrant particles. Since the new theory is highly successful, we will be forced to admit re-entrance. The claim that re-entrance is 'conceptually impossible' will thus be refuted by the creation of a physical theory which in outline is today conceptually possible and which has essential recourse to re-entrance.

III. An Implication

The conceptual impossibility of re-entrance has often been argued for by claiming that its admission would require some sort of change in the 'meaning' of the time-concept. That is, it is claimed that re-entrance is inconceivable relative to the time concept that 'we have', according to which objects cannot go backwards in time. I will here simply point out that this claim is often based on a confusion between the concept of re-entrance and that of time's 'direction'. Since I hope to treat of this elsewhere, I will not dwell on this point. I will however argue that the admission of re-entrance does not require a change in the meaning of our time-concepts.

Let us note that (P) taken alone, is not a claim about time.¹ However, if (P) is considered in conjunction with R'_1 — FR'_1 , that is, if (P) is incorporated as a framework principle of microphysics, then a new type of fact about the world is postulated, namely re-entrance. With respect to pre-Feynman physics, it is not the case that elementary particles *could not* undergo re-entrance, they simply *did not* do so. This situation should be compared with the case of special relativity. In that theory it is axiomatic that a material particle of positive mass cannot have a velocity greater than that of a photon.² In this sense (relative to the postulates of the special theory) superluminous velocities are conceptually impossible. Since comparable axiomatic restrictions against re-entrance do not occur in the present-day physics of elementary particles,³ the denial of re-entrance should be formulated as follows: There is today no law, hence no admitted process, which permits re-entrance in any but a fully conventionalistic manner. But this same

¹ In an adequate theory of the interrelation between space-time and charge symmetries this might be false, for the particle-antiparticle distinction would be connected with descriptions of possible space-time motions. This would, however, only support the remainder of my argument.

² A. Grunbaum, *Philosophical Problems of Space and Time* (New York, Knopf, 1963), ch. 12.

³ The mention here of elementary particles is of importance, since it may be the case that physical systems of greater complexity cannot be re-entrant.

absence of axiomatic restrictions shows that future extensions of particle theory (as outlined above) might admit *both* re-entrant and non-re-entrant processes, and that such an admission might best be viewed as the discovery of a new class of physical processes rather than as a change in the meaning of fundamental concepts.¹ We see then that the conceptual possibility of re-entrance is a question to be decided not by conceptual analysis but by physics.

GEORGE BERGER
Columbia University
December 2, 1967



HUMAN AND MACHINE LOGIC: A REJOINDER

We can imagine a human operator playing a game of one-upmanship against a programmed computer. If the program is F_n , the human operator can print the theorem G_n , which the programmed computer, or, if you prefer, the program, would never print, if it is consistent. This is true for each whole number n , but the victory is a hollow one since a second computer, loaded with program C , could put the human operator out of a job. . . . It is useless for the 'mentalist' to argue that any given program can always be improved, since the process for improving programs can presumably be programmed also; certainly this can be done if the mentalist describes how the improvement is to be made. If he does give such a description, then he has not made a case.²

Good, together with Turing and Benacerraf,³ argues against 'mentalism' by denying that there could be any peculiarly mental powers, since if there were, they could be described, and if they could be described, a computer could be programmed to simulate them. I am not convinced that to be is to be describable in computer language, but I shall not discuss that here: for it seems to me that even if this step were valid, the whole argument would work against only one form of mentalism, and would completely fail to establish mechanism. For if mechanism is true, a complete specification of the mental mechanism of each human being can in principle be given. But once given, it proves inadequate, in that it cannot produce as true the Gödelian formula G , which a human being can see to be true. This is no 'hollow victory' for the human being, since *this* was the specification the mechanist put forward. It is of course true that a second computer, that is, *another* computer, could do as well as a human operator on *this* test: but that is not relevant, for it was not that computer that was supposed

¹ These remarks are intended in support of Hilary Putnam's argument in his well known article 'It Ain't Necessarily So', *Journal of Philosophy*, 59 (1962). See especially his insistence that necessity is always 'relative to a body of knowledge'.

² I. J. Good, 'Human and Machine Logic', *This Journal*, 18 (1967) 145-6.

³ A. M. Turing, 'Computing Machinery and Intelligence', *Mind*, 59 (1950), 444-5, reprinted in James R. Newman, *The World of Mathematics*, iv (New York, 1956) 2110; and in Alan Ross Anderson, *Minds and Machines* (Englewood Cliffs, N.J., 1964), pp. 15-16. Paul Benacerraf, 'God, the Devil and Gödel', *The Monist*, 51 (1967), 22-23.

to match the human being. And if the mechanist now makes it relevant, by shifting his ground and saying that it is the second computer that matches the human being, then there is *another* Gödelian formula which the second computer cannot produce as true but which a human being can. In my original paper I took the argument as far as ω .¹ Good takes a further step to $\omega + 1$, and is all set to

on to somewhere between ω^ω and τ (if it exists): but the principle remains the same; the human being can always 'trump' any particular machine, and the mechanist can always produce another machine that will not fail the test which stumped the first machine. If the game played between the mentalist and the mechanist is the game of one-upmanship, the result must be a draw. But this is not the game I proposed. I proposed a matching game, not a one-upman game: the question at issue was not superiority—is the mind better than the machine, or *vice versa*?—but equality—is this, specified, machine equal, in respect of intellectual powers, to this particular person? Good thinks the mentalist must be a sort of Balliol man, and adduces arguments to show that his effortless superiority is illusory. But mine is a different argument, less arrogant, depending on a subtle shift of the onus of proof. It is not a straightforward proof, starting from some acceptable premisses and leading to the conclusion that minds are better than machines, but rather it is a schema of refutation, showing how *if* the mechanist were to particularise enough to say what machine was equivalent to a named man, the mentalist can refute at least that equivalence. But since he can do it for any specified machine, he can do it for all. Not effortlessly superior, but incontrovertibly unequal. This is not enough to re-establish traditional dualism—which seems to me open to many philosophical objections: but it is enough to refute traditional materialism and its modern mechanist version.

J. R. LUCAS
Merton College Oxford

THE WHITE SHOE *QUA* HERRING IS PINK

Dr Hempel² points out that I³ had misunderstood him and that in his context a white shoe is not a red herring. But in my context⁴ it was a red herring; so in its capacity as a herring it seems to be pink. I shall now argue its redness even within Hempel's context.

Let H be a hypothesis of the form that class A is contained in class B, for example, 'all crows are black'. Let E be what I call a 'case' of H, that is, a proposition of the form 'this object is in both class A and B'. I showed by means of a clear-cut succinct example that when there is background or given knowledge G it is possible for E to undermine H. In fact, in a technical sense, the weight of

¹ 'Minds, Machines and Gödel', *Philosophy*, 36 (1961), 117; reprinted in Kenneth M. Sayre and Frederick J. Crosson, eds., *The Modeling of Mind* (Notre Dame, 1963), p. 260.

² C. G. Hempel, 'The white shoe: no red herring', this *Journal*, 18 (1967), 239-40.

³ I. J. Good, 'The white shoe is a red herring', this *Journal*, 17 (1967), 322.

⁴ I. J. Good, 'The paradox of confirmation, II', this *Journal*, 12 (1961), 63-64.

evidence concerning H provided by E given G , denoted by $W(H:E|G)$, can be numerically measurable and negative. Hempel's thesis however is unaffected unless we can show that $W(H:E)$ can be negative, where $W(H:E)$ is in no sense an abbreviation for $W(H:E|G)$ at any rate if G has any empirical content.

Since the propositions H and E would be meaningless in the absence of empirical knowledge, it is difficult to decide whether $W(H:E)$ is necessarily positive. The closest I can get to giving $W(H:E)$ a practical significance is to imagine an infinitely intelligent newborn baby having built-in neural circuits enabling him to deal with formal logic, English syntax, and subjective probability. He might now argue, after defining a crow in detail, that it is initially extremely unlikely that there are any crows, and therefore that it is extremely likely that all crows are black, that is, that H is true. 'On the other hand', he goes on to argue, 'if there are crows, then there is a reasonable chance that they are of a variety of colours. Therefore, if I were to discover that even a black crow exists I would consider H to be less probable than it was initially.'

I conclude from this that the herring is a fairly deep shade of pink.

I. J. GOOD
Virginia Polytechnic Institute

Reviews

Karl R. Popper: *Conjectures and Refutations*. (London: Routledge and Kegan Paul, 1965 (1963), 55s.) Pp. xiii+417.

1. *Some Main Themes*

This volume of essays by Sir Karl Popper is a delight to read. Bold in style and thought, replete with choice philosophical, historical, and political insights, brimming with attacks on the status quo, it represents Popper's thought from roughly the Second World War to the present time. It can be read with great profit by Popper fans and foes alike, as well as by the uninitiated. The common theme of the essays, as suggested by the title, is that knowledge progresses by a method of conjecture and refutation. A hypothesis is proposed and subjected to scrutiny by attempts to show that it is false. If it survives such tests it is provisionally accepted, although its truth can never be known.

How does a scientist develop a theory? Not, Popper answers, simply by observing nature:

Those among you who hold the opposite view and who believe that scientific theories are the result of observations, I challenge to start observing here and now and to give me the scientific results of your observations. You may say that this is unfair, and that there is nothing very remarkable to observe here and now. But even if you go on to the end of your lives, notebook in hand, writing down everything you observe, and if you finally bequeath this important notebook to the Royal Society, asking them to make science out of it, then the Royal Society might preserve it as a curiosity, but decidedly not as a source of knowledge. It might be lost perhaps in some cellar of the British Museum (which as you may know cannot afford to catalogue most of its treasures) but more likely it will end up on a rubbish heap (p. 128).

Defending this point, Popper appeals to historical examples, claiming that great theories such as Newton's dynamics or Kepler's laws 'were not the result of observations' (p. 188). The scientist begins not with observations but with a problem or set of problems (p. 155). In trying to solve the problem he proposes a hypothesis or theory. This is a pure conjecture, on his part, that is not derived from observations. Theories are 'free creations of our own minds, the result of an almost poetic intuition . . .' (p. 192). The role of observations is not in generating a theory, but in testing it, criticising it, trying to refute it. Observations can never verify theories, since theories are or contain general statements which no finite number of observation statements imply. For this reason Popper rejects Verificationism, which he construes as requiring theories to be deducible from observation statements. Those who preach Verificationism do so as a method for distinguishing what is meaningful from what is not, another idea Popper rejects. Falsifiability, not verifiability, is the 'criterion of demarcation'. Moreover,

it is a criterion separating not sense from nonsense, but science from metaphysics. Falsifiability admits of degrees, so there is no sharp division between science and metaphysics, only differences in degree. Indeed most scientific theories originated in myths which were largely untestable, and later they developed testable components. It would be strange to say that such theories 'are nonsensical gibberish in one stage of their development, and then suddenly become good sense in another' (p. 257).

Science, for Popper, progresses from problem to problem. He compares truth to a mountain peak hidden among the clouds. The climber may never know when he has found it, though he may be getting closer and closer, and will know if he has not reached it (p. 226). By a series of conjectures and attempted refutations, the scientist can hope to progress toward the summit. There are three requirements for progress in this ascent (pp. 241-2). First, the theories proposed should proceed from some simple, new, and powerful unifying idea. Second, they should have testable consequences in addition to those they were designed to explain. Third, they should pass some new and severe tests.

Given two theories, T_1 and T_2 , we may be able to compare the degree to which they approximate the truth. The truth-content of a theory T is the class of true logical consequences of T ; its falsity-content is the class of its false logical consequences. If the truth-content and falsity-content of T_1 and T_2 are comparable, we can say that T_2 more closely approximates the truth than T_1 , or, using Popper's term, that T_2 has a greater degree of verisimilitude than T_1 , if the truth-content but not the falsity-content of T_2 is greater than that of T_1 and the falsity-content of T_1 , but not its truth-content, is greater than that of T_2 (p. 233).

So much for a brief outline of a few of the book's main, unifying themes. In what follows I propose to take Popper's advice seriously, and indicate what I think are some of the problems his view faces.

2. *The Growth of Science*

Has Popper adequately described the growth of science? What he says gives the impression that if we observe what the scientist does, at least the good scientist, we will see him proposing a theory in response to a problem or set of problems and then subjecting this theory to severe tests; if the tests falsify his theory, a new theory will be invented and this in turn tested; and so it goes. On this view, the history of science, at least good science, is a history of conjectures and attempts at refutation. Progress in science is achieved only by proposing and testing theories.

No doubt Popper has characterised important elements in scientific progress. But do all elements fit neatly into his scheme? Take Millikan's oil drop experiment. In performing the experiment, that is, in setting up the apparatus consisting of a chamber with oppositely charged plates through which drops of oil could fall and be illuminated, and in observing the results through his short focus telescope, Millikan was not proposing a theory. Was he then testing one? Admittedly, as Millikan himself says, one of his aims was to test the theory that electrical charges are exact multiples of an elementary electrical charge. But another aim, stated by Millikan, was 'to make an exact determination of the value of the elementary electrical charge which is free from all questionable

theoretical assumptions.¹ True, previous estimates of this value had been made, but these were based on imprecise methods, and anyway, Millikan was not testing *these* hypotheses. There was no hypothesis regarding the value of this quantity that was being tested by the experiment. When Millikan determined the elementary charge to be 4.93×10^{-10} esu, other tests of this value could have been made, and indeed were. (Later corrections put the value at 4.8025×10^{-10} esu.) But one of the important things Millikan was doing was neither testing a previously proposed hypothesis, nor testing one which the experiment was supposed to generate. Although he hoped to arrive at a certain result by means of the experiment (a result that Popper might call a theory or conjecture), the experiment itself could not be described simply or solely as the proposing of a theory, as the testing of one, or even as both. Despite this, Millikan's experiment represented important progress in physics.

Again consider what Maxwell did in his famous 1885 paper, 'On Faraday's Lines of Force'.² His aim was neither to propose nor to test a hypothesis regarding the nature of electrical phenomena. It was rather to offer a 'simplification and reduction of the results of previous investigation to a form in which the mind can grasp them'.³ This Maxwell did by elaborating a physical analogy between the electric field and tubes containing an imaginary, incompressible fluid. For example, he showed how electrostatic force is analogous to the velocity of his fluid in that both quantities vary inversely as the square of the distance from respective sources, whereas the potential of the electric field at a given point is analogous to the pressure of the fluid at a given point since both quantities vary inversely as the first power of the distance. In constructing this analogy he was neither testing a theory nor proposing one. Maxwell writes:

By the method which I adopt, I hope to render it evident that I am not attempting to establish any physical theory of a science in which I have hardly made a single experiment and that the limit of my design is to show how, by a strict application of the ideas and methods of Faraday, the connexion of the very different orders of phenomena which he has discovered may be clearly placed before the mathematical mind.⁴

Maxwell later developed his electromagnetic theory without invoking this analogy, but it would be difficult to maintain that the analogy made no positive contribution to the growth of field theory.

I have mentioned experimentation and the construction of analogies neither of which fall under the rubric of proposing or testing theories, yet both of which can contribute to the growth of science. Are there other such activities? Suppose a scientist attempts to reformulate in mathematical terms previously proposed hypotheses which were expressed qualitatively, as Clausius did with the first and second laws of thermodynamics; or suppose he introduces and defines new concepts, as Gibbs did in the case of an ensemble in statistical mechanics; or suppose he attempts to redefine a concept, as Mach did with mass; or suppose he applies a theory to special, idealised systems and works out the results, as

¹ Reprinted in *Great Experiments in Physics*, ed. Morris H. Shamos. (Holt, Rinehart and Winston, New York, 1962), p. 242.

² *The Scientific Papers of James Clerk Maxwell*, ed. W. D. Niven (Dover Publications, New York, 1965), I, 155-229.

³ *Ibid.* p. 155.

⁴ *Ibid.* pp. 157-8.

Clausius did when he applied the laws of thermodynamics to a perfect gas; or suppose he presents illustrations of certain principles and definitions. Can these activities be described simply as the proposing or testing of theories?

In some cases, such as defining or presenting illustrations, the answer seems to be clearly No. In others, such as providing a mathematical reformulation or applying the results to idealised systems, it is less clear, though, again, I should be inclined to answer No. When Popper speaks of conjecturing he seems to have in mind the proposing of new theories, not the reformulation of old ones; moreover, a scientist could reformulate a principle and even apply it to special systems without proposing it, that is, without putting it forth as a conjecture, and, of course, without testing it. In any event, the book does not offer as much clarification on this score as one might like. Popper does mention (p. 221) the activity of axiomatising theories, and he recognises the importance of making deductions from a theory in order to test it. But some of the activities I have mentioned here or earlier (for example, drawing analogies, performing experiments, citing illustrations) would not be involved in axiomatising a theory or in generating consequences from it, and none need be so involved; that is, each can be, and often is, done independently of an attempt to axiomatise a theory or derive consequences from it. As already reported, Popper does cite three requirements theories must satisfy if growth of knowledge is to take place: theories must proceed from a simple, unifying idea; they must yield new and testable consequences; they must pass some new and severe tests. But these requirements say nothing about the sorts of activities mentioned above.

Undoubtedly, Popper would not deny the existence of these activities. Perhaps he will say that they do not promote scientific progress. If so, there is no argument given to show this, and, I think, it would take Herculean efforts to produce one. More likely, he will say that they promote progress only in so far as they lead to the proposing or testing of conjectures: they can be useful stepping stones, but of little or no scientific value if considered by themselves.¹ Unless he defines progress so as to make this claim true, I cannot accept it. Someone like Maxwell, reformulating in mathematical terms previously expressed conjectures of kinetic theory by pursuing what he called 'an important physical analogy' between gaseous systems and those composed of small elastic spheres, promoted at least as much progress in physics as someone like Boyle, who simply conjectured that gases might be composed of corpuscles but did not work this out in any detail, or as someone like Stern, who by developing molecular beam methods subjected some of the principles of the kinetic theory to experimental test. Moreover, Maxwell achieved progress not simply in virtue of the fact that he reformulated the theory so as to make it more testable, more criticisable. If we must focus on any one thing, I should say he achieved progress in virtue of the fact that by pursuing his analogy, and by dealing in a mathematical way with issues concerning the mean free path of molecules, velocity distributions, interpretation of viscosity in molecular terms, and so forth, he helped promote greater scientific understanding of the molecular nature of gases. More generally,

¹ This is what Popper explicitly says about the activity of axiomatising theories. For example, he writes: 'Consequently, there is little merit in formalising and elaborating a deductive non-conventional system beyond the requirements of the task of criticising and testing it, and of comparing it critically with competitors' (p. 221).

I would speak of the promotion of scientific understanding as at least a sufficient condition for progress in science. Inventing and testing new conjectures can promote such understanding, but so can each of the other activities mentioned, and not simply by serving as stepping stones to new conjectures and tests.

3. *The Role of Observation*

I turn now to a second issue arising from Popper's account of the growth of scientific knowledge. Popper speaks of theories as '*free* creations of our own minds, the result of an almost poetic intuition' (p. 192), as 'our own inventions, our own ideas; they are not forced upon us, but are our self-made instruments of thought' (p. 117). More particularly, they are not the results of observation, whose role consists only (or mainly) in testing them.¹ From these passages one might infer that Popper is claiming that theories develop completely independently of one another; they have no historical roots in previous theories. Yet this he explicitly denies in other passages, when, for example, he claims that 'the advance in knowledge consists, mainly, in the modification of earlier knowledge' (p. 28), or that 'most of our scientific theories originate in myths'. So, I think, the best way of construing his position is this: Theories are free creations in the sense that they are not forced upon us either by observation or by previous theories. This means that the mere existence of a previous theory does not necessitate the adoption of any particular theory, and that no set of statements reporting what has been observed logically implies any theory. With this I think I can agree. It is Popper's conclusion, that the role of observation consists solely (or mainly) in testing theories, that I cannot accept.

Observations can be relevant also in suggesting a theory, even though they do not logically entail it. I am not saying that observations by themselves, independently of any other knowledge, suggest theories. But they can play a part, whose importance can vary. Consider scientific laws. Some, to be sure, were developed largely independently of observations: Maxwell derived his distribution law from theoretical assumptions of the kinetic theory. But in other cases observation played a much more prominent role.

Consider the conservation of strangeness in contemporary nuclear physics. Gell-Mann and Nishijima observed that certain nuclear reactions, allowed by hitherto known conservation laws, do not occur in accelerator experiments. If a high-energy negative K-meson collides with a proton they observed that the interaction yields pi-mesons, K-mesons, lambda particles, and sigma particles—but only in certain combinations. For example, there is never a proton and a negative pi-meson, although this is not precluded by other principles such as conservation of charge, energy, and momentum. These experimental results suggested to Gell-Mann and Nishijima that another conservation principle was at work.

¹ At one point (p. 28) Popper suggests that this is their main role. At the other points he says it is their only one, for example: 'In the development of science observations and experiments play only the role of critical arguments. And they play this role alongside other, non-observational arguments. It is an important role; but the significance of observations and experiments depends *entirely* upon the question whether or not they may be used to *criticize theories*' (p. 152). Popper does not deny that any particular theory may be preceded by observations which the theory is designed to explain (p. 47). But, he insists, the role of observations is to be construed as that of testing theories, not generating them.

They proposed that each particle be assigned what they called a strangeness number and that in any strong interaction the total strangeness before must equal the total strangeness after. Admittedly, this principle was developed by them against a background of other principles. But the observations these physicists made played an important role, indeed the crucial one, in the development of the principle of the conservation of strangeness. Statements describing these observations do not entail this principle, nor were the observations reported in order to test the principle. The observations these physicists made, in view of the accepted background of theory, suggested that another conservation principle was operative.

How should this concept of suggestion be described? Here we might consider some ideas of the late N. R. Hanson.¹ Hanson, following Peirce, speaks of *retroduction* as characterising the logic of scientific discovery. On his view, the scientist begins with some observed phenomena which are surprising or puzzling, and by studying these hits upon a theory which organises them into an 'intelligible, systematic, conceptual pattern'. He characterises the (retroductive) inference from data to theory as follows:

Some surprising phenomenon P is observed.
P would be explicable as a matter of course if H were true.
Hence there is reason to think that H is true.²

Hanson rejects the hypothetico-deductive account of theories on the ground that, contrary to what the latter suggests, scientists do not begin with hypotheses and then deduce the observed data. They begin with the observed data and seek hypotheses to explain them. He also rejects the inductivist account on the ground that the scientist does not arrive at his theory by making inductive inferences from the data. The mode of inference actually employed is retroduction. On Hanson's account, to speak of observations as suggesting a theory is to speak of a retroductive inference from the observations to the theory.

Despite the fact that they both reject inductivism, Popper and Hanson are at opposite poles. Hanson sees the scientist as beginning with observations and making a retroductive inference to a theory. Popper sees the scientist as freely inventing a theory (against a background of other theories) and making observations only to test it. I think both accounts are too extreme, though both have important elements of truth. Popper is right in stressing, and Hanson mistaken in neglecting, the background of theory which the scientist frequently has to begin with. Hanson is right in stressing, and Popper mistaken in neglecting, the role observations often play in developing a theory. Contrary to what Hanson suggests, the supplier of a theory does not always begin simply by contemplating observed surprising phenomena. He may begin by considering something neither surprising, nor a phenomenon, nor even observed. Niels Bohr, in developing his theory of the atom, began by considering the structure of the hydrogen atom postulated by Rutherford's model and the various problems such a model posed. Moreover, he considered this not in isolation but against a background of theory: classical mechanics, classical electrodynamics, and Planck's theory of radiation. On the other hand, contrary to what Popper suggests, the scientist often begins

¹ N. R. Hanson, *Patterns of Discovery* (Cambridge University Press, 1958), ch. 4.

² *Ibid.* p. 86.

by considering, among other things, observations, which, in the light of background theories and the problems these face, may suggest the plausibility of a new theory to explain the observations. I do not claim this is always the procedure of the scientist, but it frequently is. So, if we do not take Hanson's retroductive inference as characterising all inference to theories, and if we suitably modify it, I think we can say that it characterises the relationship between observation and theory in a number of important cases. The modified version might be this:

Against such and such a theoretical background, some surprising phenomenon P is observed.

P would be explicable as a matter of course if H were true (which may require important changes in the theoretical background).

Hence there is reason to think that H is true.

For example:

Against the background of nuclear theory and conservation laws, Gell-Mann and Nishijima observed that certain interactions permitted by this theory do not occur.

This would be explicable as a matter of course if a further conservation principle of such and such a sort were operative.

Hence there is reason to think that such a conservation principle is operative.

To speak of the observations of Gell-Mann and Nishijima as suggesting the conservation of strangeness is to speak of an inference of this type.

4. *Testing Theories*

Throughout the book, Popper emphasises that theories are guesses which cannot be verified, justified, established, or known to be true. They can only be tested and disproved. Here are some typical passages:

So my answer to the questions 'How do you know? What is the source or basis of your assertion? What observations have led you to it?' would be: 'I do *not* know: my assertion was merely a guess' (p. 27).

[it is a] false idea . . . that we must *justify* our knowledge, or our theories, by *positive* reasons, that is, by reasons capable of establishing them, or at least of making them highly probable (p. 29).

. . . nothing can be justified or proved (outside of mathematics and logic) (p. 51).

. . . scientific theories . . . are *genuine conjectures*—highly informative guesses about the world which although not verifiable (i.e. capable of being shown to be true) can be submitted to severe tests (p. 115).

. . . we do not try to prove or to *verify* them [theories], but we test them by trying to *disprove* or falsify them, to *refute* them (p. 192).

. . . we can never give positive reasons which justify a belief that a theory is true (p. 226).

One's immediate reaction to these claims might be this: Surely we ought to demand more of science than mere guesses which are not verified, justified, or established, but can only be tested and disproved. Otherwise, why pay more attention to the scientist than to the crystal ball gazer whose guesses can also be

tested and disproved? To the extent that a proposition is a mere guess which has not been justified, it is unreasonable to accept it and act upon it. However, those who might react in this way to the claims quoted above must also remember that Popper requires a theory to 'pass some new, and severe, tests' before provisional acceptance (p. 242). Such tests must be ones designed to try to disprove the theory. Suppose a given theory passes all such severe tests with flying colours: every time we try to disprove it we fail to do so; the theory always yields new predictions 'corroborated by experimental evidence' (p. 243). Our second reaction to the quotations above might then be this: Since our theory stands up to severe tests, it is verified, justified, established. There are positive reasons in its favour. Although once it may have been a guess, it is no longer one. I think this second reaction is a reasonable one, for to subject a theory to such tests which it then passes is (at least part of) what it means to verify a theory, justify it, establish it, or have positive reasons in its favour. The mere fact that in testing our theory we have tried to disprove it by using controlled experimentation instead of trying to 'prove' it by haphazard and indiscriminate observation does not make the results of our tests any less positive or justificatory. On the contrary, it makes them more so.

What has gone wrong, I think, can be explained by reference to Popper's treatment of early Positivism. The early Positivist accepts the following argument:

Early Positivist Argument

- (1) A theory is verifiable if and only if it is deducible from observation statements.
- (2) Scientific theories are deducible from observation statements.
- (3) Hence, scientific theories are verifiable.

Popper rejects the conclusion of the argument because he rejects premise (2). He would propose the following argument:

Popper's Argument

- (1) A theory is verifiable if and only if it is deducible from observation statements.
- (2') Scientific theories are not deducible from observation statements.
- (3') Hence, scientific theories are not verifiable. Therefore, they cannot be established or known to be true, so they are guesses.

Popper has rejected the thesis of the early Positivists that scientific theories are deducible from observation statements, but has retained their doctrine that a theory is verifiable if and only if it is deducible from observation statements. He has used the latter as a basis for his startling claim about knowledge: since scientific theories cannot be deduced from observation statements, they cannot be verified, and therefore cannot be known to be true; so they must be guesses. The less startling course of action, and, I would suggest, the more reasonable one, would be to reject not only the Positivist premise (2) but premise (1) as well, since the latter provides a patently incorrect account of verification.¹

¹ Indeed, in later writings Positivists introduced notions of verifiability that did not require deducibility from observation statements. See, for example, A. J. Ayer,

It might be asked: why all this fuss over the word 'verification'? Give this word to the early Positivist, and let us use a different one, for example, 'corroboration', a term introduced by Popper himself. We will say that although theories cannot be verified they can be corroborated by passing severe tests.

Fine. The problem is that in making many of his claims by continuing to use 'verification' in the early Positivist sense, Popper's general theory of knowledge has been affected. From the fact that theories are not verifiable (in the early Positivist sense) he concludes that they cannot be established or known to be true, and must be guesses. But this conclusion does not follow unless, of course, he is also using terms like 'establish', 'know', and 'guess' in exotic senses that early Positivists might have given them. And if he is, then his epistemological claims are not very startling at all.

One final point. Suppose Popper insists that he is using terms like 'establish' and 'know' in standard ways, and that theories cannot be established or known to be true. Then, contrary to what he suggests, I think it also follows from his position that theories cannot be tested. According to Popper, observations can only falsify but cannot verify a theory: '*Only the falsity of a theory can be inferred from empirical evidence, and this inference is a purely deductive one*' (p. 55). The doctrine appears to be that we make observations which are described by means of so-called observation statements. If the latter are logically incompatible with the assumptions or consequences of the theory, the theory is falsified. Popper speaks of *basic statements* as ones which 'state (truly or falsely) the existence of observable facts (occurrences) within some sufficiently narrow spatio-temporal region' (p. 386).¹ In the light of theories, observations are made and basic statements are used to describe these observations. These basic statements can be logically incompatible with the theory, and hence are used to test it. However, Popper emphasises, basic statements themselves are not verifiable because they contain general terms denoting items exhibiting law-like behaviour (p. 387). Basic statements are implicitly general, and so are also conjectures.

Now I take it that to falsify a theory is to show or establish that it is false; it involves coming to know that it is false, and not simply claiming that it is false. If basic statements are to falsify a theory they must not merely be asserted, they must be established, they must be known to be true. This cannot be, for Popper. The most we can do is conjecture that basic statements are true and hence (if these are incompatible with the theory) that the theory is false. But this is not to falsify it. It is not to test the theory, if, in principle, tests cannot be effective. So if Popper insists that he is using terms like 'establish' and 'know' in standard ways, and if he insists that neither general theories nor basic statements can be established or known to be true, then, I think, he must admit not only the impossibility of verifying theories but also the impossibility of falsifying them, in short, the impossibility of testing them.

Language, Truth, and Logic (Victor Gollancz, London, 1937; 2nd ed., 1946); or even later, Rudolf Carnap, 'The Methodological Character of Theoretical Concepts', *Minnesota Studies in the Philosophy of Science*, Vol. I ed. H. Feigl and M. Scriven (University of Minnesota Press, Minneapolis, 1956).

¹ He goes on to say that the negation of a basic statement is not basic, but that the conjunction of two basic statements is basic, provided it is logically consistent, and he indicates how even further basic statements can be constructed from these.

In a footnote Popper acknowledges that this objection might be made, and he suggests a reply.¹ The objection, he says, mixes two different levels of analysis. On one level there is a logical asymmetry: a theory can be formally falsified by any number of basic statements but cannot be verified by any number of them. On another level, he admits, 'we may hesitate to accept any statement, even the simplest observation statement'. Yet, he concludes, the latter fact, although important, 'does not affect the fundamental asymmetry'.

I agree with Popper that there is an asymmetry, that although no theory is logically deducible from basic statements it may be logically incompatible with them. But this does not diminish the present difficulty. I am not simply pointing out that any statement, including basic ones, can be subjected to further tests or that 'we may hesitate to accept any statement, even the simplest observation statement'. My point is that if basic statements are to refute a theory they cannot merely be conjectures; they must be known to be true. Popper says that basic statements cannot be known to be true, they can only be conjectures. It follows that, for Popper, theories cannot be refuted.

This yields an even more bizarre consequence. Popper says that if we subject a theory to severe tests and it is not refuted then we can accept it, albeit tentatively (p. 51). But if we can never refute a theory, no matter how severe the tests, then we can accept any theory we like, and do so 'untentatively.'

PETER ACHINSTEIN

The Johns Hopkins University

Stephan Körner: *Experience and Theory* (London: Routledge and Kegan Paul, 1966, 45s.). Pp. xiii+250.

Professor Körner's book is primarily, as its sub-title says, an essay on the philosophy of science. But not exclusively: it starts with a general account of empirical concepts, two chapters are concerned mainly with the philosophy of mathematics, and at the end it touches upon mental phenomena, freedom of choice, and religion.

As any reader of Körner's previous books would expect, this one too is learned, closely-reasoned, ingenious, and original. Its central thesis is that 'empirical and theoretical discourse are logically disconnected'. The concepts of theoretical science differ in kind from those of ordinary thought. Theoretical concepts, and the logic they obey, are 'idealisations' of their empirical counterparts; consequently there can be no strict deductive relations between the two spheres. A theory, together with observations, never *entails* the prediction of further observable events; nor can a theory be conclusively falsified by observations.

Some of the other theses about similar disconnections may matter less to the philosopher of science. Mental phenomena are 'intentional' in Brentano's sense, and are therefore neither reducible to physical phenomena nor amenable to the methods of the natural sciences. Free choices—causally independent of preceding conditions but causally effective in bringing about later events—

¹ Page 41, note 8. See also Popper, *The Logic of Scientific Discovery* (Hutchinson, New York, 1961), pp. 41-42.

cannot be ruled out by any physical theory. And some theological notions, like those of a perfect being and of creation from nothing, but not that of miracles, are compatible with all scientific theories.

The arguments that support the theses of this second group are comparatively brief and non-technical, in contrast to the argument for the central thesis, with its elaborations, which occupies most of the book, and makes considerable use of logical symbols; only the serious student of the philosophy of science will make much of it. On the philosophy of mathematics in particular Körner writes allusively, referring to the ideas and results of much recent work, and while I, for one, want to question parts of what he says here, I should have to dig a long way down into the conceptual foundations before any fruitful discussion could begin.

However, the main outline of Körner's argument is clear enough, and can be discussed at a less technical level. His central thesis rests ultimately on a view about 'empirical discourse'. Empirical classes are inexact: for example, while some things can be confidently classified as green, and others as not green, there are also 'neutral candidates', things such that it is a matter for free decision whether to count them as green or not. For this and related reasons, some statements in empirical discourse are neither true nor false but 'neutral': they may, by a free choice, be made either true or false. From this Körner infers that they obey a different logic from the ordinary one which admits only truth and falsehood and assumes that all classes are exact. It is this that, in the end, necessitates the logical disconnection between theory and experience. Theory works with exact classes and is embedded in classical logic. Within a theory conclusions are deduced from premisses, laws entail predictions, and so on, but there is a gap, which no deduction can bridge, between these theoretical conclusions and predictions and the corresponding observation statements.

In this Körner is going far beyond the claim, which would be universally conceded, that scientific theories often differ from the corresponding observational data in their precision, abstractness, and simplicity. On the other hand, he admits that the working scientist may ignore the disconnection: we 'identify' a proposition of theoretical physics, say, with a perceptual one, that is, we close the logical gap by treating them as if they were identical, though they are not. Körner's thesis bears not upon scientific practice, but upon the philosophical understanding of it and of its relations to other pursuits.

It seems to me, however, that the inexactness of empirical discourse is too slight a foundation to carry the weight of the superstructure Körner builds on it. For one thing, he admits that ordinary 'classical' logic and the 'modified two-valued logic' appropriate to empirical discourse largely coincide. They must, since some predicates even of empirical discourse are exact, and once its neutral propositions have, by decision, been made true or false they must obey laws which agree with the classical ones. But, secondly, the differences which Körner still finds between the two logics are open to question. They lie not in the valid forms (which are the same on the two logics) but in their use in deductions. Körner says that we can deduce q_0 from p_0 and (if p_0 then q_0) only if these premisses 'are either true or, in case that they are neutral, evaluated as true', hence that 'in applying *modus ponens* we ignore the neutrality of propositions', and hence that 'the classical two-valued logic as an instrument of deduction . . .

presupposes that neutral propositions are treated *as if* they were true, and inexact predicates *as if* they were exact' (pp. 46-47). Unless I have failed to understand Körner's line of thought here, it incorporates an unjustifiable slide from the fact that a valid deductive argument guarantees the truth of its conclusion only if its premisses are true to the false claim that in using any such argument we presuppose that its premisses are true. This claim must be false, for it would rule out the use of *reductio ad absurdum* arguments even within classical logic. This discussion (on pages 45 to 47) is quite brief, but the central argument of the whole book rests on it, and in what follows Körner appeals repeatedly to its conclusion. Will it stand the very close scrutiny that it undoubtedly demands?

Thirdly, even if Körner could thus establish the disconnectedness of the two forms of *discourse*—theoretical and empirical—this would not open any like gap between scientific theory and *what is observed*. The inexactness of predicates and classes is at most a contingent feature of the *language* in which we record observations; from this nothing could follow, for example, about whether a set of observations can ever conclusively falsify a theory. (Often they cannot, but this is for other reasons which Körner is not alone in mentioning, not because there is a *general* logical hiatus between experience and theory.) Similarly, Körner cannot be right in analysing empirical continuity—for example in the colour continuum—which is a perceptual matter, in terms of the inexactness of classes, which is a linguistic one.

In discussing mental phenomena, Körner rightly stresses their 'intentionality', the fact that a mental state like desire or belief involves an 'intentional object', which does not 'exist outside the mind'. But this significant but puzzling feature seems to me to call for a more searching analysis. His account again seems too linguistic, too much confined to the fact that statements about desires, etc., obey a logic which is not purely extensional. By itself, this will not show what he wants, that mental *states* cannot be reduced to physical ones. Körner goes on to argue cogently against behaviourism, but granting that intentional mental phenomena are real and are known introspectively, it does not follow that the methods by which they are studied must differ radically from those of the natural sciences. Körner's case against physicalism in general, as opposed to behaviourism, is far from complete. He assumes that we could identify mental states with physical ones only by asserting the obvious falsehood that mental terms are *synonymous* with certain physical terms. This ignores the currently important version of physicalism, the 'contingent identity thesis' being advanced by some philosophers in the United States and Australia. Nor would Körner's arguments tell against a double-aspect theory of the mind-body relation.

This affects what he says about choice. He develops an ingenious account of how causally independent but effective choices could be harmonised with a physical theory that left room for them. But his vital claim is that *every* physical theory must leave room for them. This rests partly on the central thesis, which I have already queried, about a logical gap between theory and experience, and partly on the view that physical theory not only *must* ignore mental events (because they are intentional) but *can* ignore them, because it is only meant to apply to, and only tested in, choice-free contexts. But if mental features were contingently identical with physical ones, or even if the mental and physical were two aspects of a single process, any process which included a choice would

also call for detailed physical description, and if there were uncaused choices a physical theory which did not provide for uncaused events would be false.

J. L. MACKIE
University College,
Oxford.

Imre Lakatos, Ed: *Problems in the Philosophy of Mathematics* (Amsterdam: North-Holland, 1967, 72s.). Pp. xv+241.

This is the first volume of the Proceedings of the International Colloquium in the Philosophy of Science held in London in 1965, and contains revised versions of the nine papers presented in the Philosophy of Mathematics Section, together with comments by participants in the discussions, and replies. (The papers on Inductive Logic and Philosophy of Science will be published in two separate volumes.) In a short review it is not possible to give much more than an outline of the contents.

Á. Szabó, in 'Greek Dialectic and Euclid's Axiomatics', discusses the replacement of empirical and illustrative proofs by the axiomatic method, claiming that Greek mathematicians were influenced by Eleatic dialectic. For example, unproved axioms correspond to agreed bases for philosophical debate, and indirect (reductio) proofs correspond to a method of disproving hypotheses in debate. The discussion papers (by Kneale, Kalmár, Robinson, Lucas, Bernays, Whitrow and Popper) raise such questions as: whether the Eleatic philosophers had not themselves learned from earlier mathematicians; whether Plato influenced the axiomatisation of mathematics or was merely, like the mathematicians, influenced by the Eleatic philosophy; whether the Greeks knew that different axiom systems might be equivalent; whether Euclid's systematisation of geometry was inspired more by cosmological than by mathematical problems; and the connection between mathematics and religious ritual.

In 'The Metaphysics of the Calculus', Abraham Robinson outlines the basic ideas of his Non-Standard Analysis, and shows how it can be taken as a formalisation of much of what was said about infinitesimals by the early developers of the Calculus (and repeated by some modern physicists). The discussion (Geach, Freudenthal, Heyting, Bar-Hillel and Bunge) raises, among others, the question of the connection between technical and metaphysical problems, without going into it very deeply.

Fred Sommers in 'On a Fregean Dogma' advocates an Aristotelian view of predication. He rejects the quantifier analysis of general statements, and claims that the predicate *mortal* can be affirmed of either the singular subject Socrates or the general subject *Men*. This theory, which leads to an interesting algorithm for testing syllogisms, requires statement negation (' $\sim$ ') to be distinguished from both the operation of denying (as opposed to affirming) a predicate, and the operation of forming the 'logical contrary' of a predicate. The discussion (Kalmár, Dummett, Lejewski, and Quine) brings out several obscurities and difficulties.

The next two sections are based on papers by A. Mostowski and P. Bernays on

recent results in set theory. The former gives a very clear summary of a number of recent independence theorems, mostly concerning strong axioms of infinity, from which he concludes that 'axiomatic set theory is hopelessly incomplete', and that there is a multitude of different notions of 'set' all equally admissible as the intuitive basis for set theory, and thus for mathematics. Bernays seems to agree with this general conclusion and illustrates it by discussing P. J. Cohen's independence proof for the continuum hypothesis, claiming however that it applies only to a *strictly formalised* axiomatisation of set theory. Mostowski's paper is discussed by Kreisel, Robinson and Kalmár, and Bernays's by Mostowski, Bar-Hillel and Suppes. Both Kreisel and Suppes claim that the possibility of a single objective foundation for mathematics is still open.

S. Körner's paper 'On the relevance of Post-Gödelian Mathematics to Philosophy' briefly surveys the history of the idea of a unique mathematical truth, and uses independence proofs as a basis for rejecting both geometrical intuition and set-theoretical intuition as means of discovering *the* mathematical truth. (He apparently overlooks the distinction between the geometrical case and the independence of the continuum hypothesis, stressed by Mostowski on p. 114.) The possibility of a Kantian argument for a unique set theory is considered and rejected. Körner's positive suggestions make use of two ideas expounded too briefly here for clarity, namely the idea of two incompatible mathematical theories being 'true in different possible worlds' (is there a 'possible world' in which there are only finitely many integers?) and the more promising idea of two incompatible mathematical propositions being 'coidentifiable within certain contexts'. Müller attempts to clarify and illustrate the latter concept in his comments, and Bar-Hillel criticises Körner's neglect of Carnap's ontology-free philosophy of mathematics.

The longest paper is G. Kreisel's 'Informal Rigour and Completeness Proofs': a mixture of lively philosophical discussion and technical expositions of his ideas for formal analysis of such intuitive, informal, ideas as logical validity and the evidence which a thinking subject has for asserting a mathematical statement. There are two appendices, on 'Axioms for explicitly definable properties' and on 'Standard and non-standard models'. Although the technical details of the paper and comments (especially Myhill's) will be of interest mainly to experts in recent set theory and intuitionistic mathematics (a class which does not include this reviewer), Kreisel's programmatic and philosophical remarks should be found generally stimulating and provocative.

L. Kalmár's paper, 'Foundations of mathematics—whither now?', revives the suggestion that mathematics should be regarded as an empirical science, which like other sciences has to be tested in practice. He suggests that further study of the implications of this view promises to be a fruitful field of foundational research. The commentators include Bernays, who claims that the peculiarity of mathematics is its reliance on *mental* experience, Bar-Hillel, who throws up his hands in horror yet again, and Lakatos, who points out that Kalmár has expressed a fairly wide-spread mood, offers an interesting diagnosis and suggests suitable medicine.

The final section is based on A. J. Easley Jr.'s paper 'Logic and Heuristic in Mathematics Curriculum Reform' which reports on the conflict in America between those who favour basing the teaching of mathematics on highly general

logical concepts, and those who suggest teaching heuristic methods and concrete examples as a basis. Easley suggests a third alternative combining the best of the 'logical' and the 'heuristic' approaches. He stresses (like Suppes) the need for further empirical research on the relative merits of different teaching schemes, and suggests that the methods of cultural anthropologists should be employed in this research. Kalmár contributes a report on teaching experiments in Hungary. Here, as in some of the previous sections, the discussion papers considerably enhance the value of the book.

Readers with appropriate interests are likely to find the bibliographies supplied by Mostowski and Easley useful. The book is certainly worth recommending to one's library, though as so often with such collections it is not clear how many individual readers will find enough in it to make it worth buying. There is certainly something for almost anyone interested in the history, philosophy, or teaching of mathematics.

A. SLOMAN
University of Sussex

Saul A. Basri: *A Deductive Theory of Space and Time*. (Amsterdam: North-Holland, 1966.) Pp. xi+163.

This book is an attempt to put the theory of space and time on a deductive foundation. It 'starts with the primitive concept of the class of living things and ends with the derivation of the equation of motion of a particle in an arbitrary gravitational field'. It is a purely macroscopic theory, based on concepts such as particles, events, clocks, measuring instruments, etc., but it is hoped by the author that it will serve as a springboard for the development of a microscopic theory.

The author's standpoint is to take current notions about classes, relationships, and sets of propositions specifying these relationships as fundamental for physics. In this point of view, the main language is assumed to be that of symbolic logic, while ordinary languages (such as English) are regarded merely as intermediaries, useful for communicating the symbolic logic language (SL), but thereafter to be dropped (as the scaffolding is taken down after the completion of a building).

In terms of this language, the author gives a detailed analysis of many interesting questions, such as time order, world lines, signal relations, clocks and time intervals, length measuring instruments, space geodesics, trajectories, etc. It would, however, be difficult to comment extensively on this analysis here, because it is so detailed and condensed in its structure.

It is the author's suggestion that the reader will be able to think directly in terms of SL, rather than in the vague and ambiguous terms that one inevitably has to use in ordinary languages. While the reviewer sympathises with those who want to remedy the many deficiencies in the common language, both obvious and subtle, he is not at all sure that to think in terms of SL is the way to do this. Since the author directly addresses himself to physicists, and since the reviewer is a physicist, it will perhaps be useful to summarise the reviewer's responses to

uch notions, if only to help orient the author's mode of presentation in the projected extension of his deductive programme to the whole of physics.

The author's basic assumptions, on which all else is founded, are given in chapter II. He begins by quite pertinently remarking that science is after all a human activity. He then goes on to say that 'our conception of the Universe is based directly upon our sensations and how we interpret them'. Thus it seems natural to the author that theory should start with the concept of a living human being as an observer and lead to the concept of the objective Universe through the analysis of the sensations of observers.

Further on, in section 2, the author takes as basic the assumption of a '*class of living humans* who have *adequately functioning sense organs*, can *communicate* with each other, and do so *honestly* and *without bias*'. The author disqualifies all 'whose senses are preconditioned to influence future sensations, as well as those who are under the influence of drugs, or are reputed to be deceitful or biased'. Who can qualify as an observer is said to be a question requiring long experience with comparison of sensations. The author is satisfied that a 'sufficient' number of such observers do exist.

The author points out that the concept of an object is very complicated. He feels that his own approach will lead to greater clarity and deeper understanding than would the direct introduction of the concept of the object.

The reviewer is in sympathy with the author's notion that the concept of the object provides a complicated and perhaps confused starting point in such thinking. As much as he realises the need to try to begin somewhere else, however, he feels that the author's basic assumptions are even more complicated and confused than is the concept of an object.

Consider only one point, for example, the concept of *sensations*. What exactly *is* a sensation? The author seems to imply that we can directly observe our sensations and identify them (whereas we find a lot of difficulty when we try to say what is meant by observing an object). Perhaps the reviewer has missed what is the real meaning of the author here. But when he tries to observe a sensation 'directly' all that he can find is an unanalysable and unspecifiable flux, in which there is nothing to be recognised or described. Such a simple sensation as what is *called* pleasure or pain reveals itself as an infinitely complex and subtly varying pattern of change, in which the variations in intensity of pleasure are indistinguishable from pain, while variations in pain give rise to pleasure, fear, anticipation, etc. A similar quality of ineffability reveals itself when the reviewer tries to specify his visual sensations, his auditory sensations, etc. As far as the reviewer can tell, the concept of a sensation is a purely verbal abstraction, almost infinitely more diffuse and shadowy than is the (admittedly unclear) concept of an object.

Then what is an 'observer'? This question (that of the nature of subject and object) is without doubt the most entangled and confused one in the whole field of human thinking. The chief difficulty is that if subject and object are assumed to be entirely different in their natures, we cannot obtain a clear account of how they are to be *related*. What usually happens is that, in effect, their relationship is explained by taking the subject (as observer) to be just another object, one among many. But then, one needs yet another subject to see that this is the case. What about this other subject? Is it too just another object? If so, we are started off on

an infinite regress with no solution. If not, we must regard some 'ultimate subject' as totally different from all objects. Some people, e.g. Bishop Berkeley, have assumed that this 'ultimate subject' is God. However, the author does not appear to be proposing this solution.

At the beginning of section IV, he suggests that sensations are correlated and that these correlated sensations are attributed to a 'subjective entity'. But the difficulty is just the one indicated above, i.e. that this 'subjective entity' is effectively treated as yet another object. Although he has been *defined* to be a subject, nevertheless the mode of definition is exactly that by which we always establish our objects, i.e. attributing correlated observable aspects to a common entity, assumed to be their source. So, in fact, the 'subjective entity' differs from an 'objective entity' only in name, and not in the actual way in which the concept is constructed, and in which it operates in the whole field of concepts. That this is actually the case can be seen yet more sharply by noting that we are inevitably led to the question: 'Who is the entity that experiences the sensations and does the correlating?' Evidently, since the 'subjective entity' is just one of his thoughts, he must be a subject of another order. But then he, too, has been established by correlations (of concepts rather than sensations). Therefore, he also is only an idea in the mind of a yet higher subject, and so on, ad infinitum (or until we reach God).

When we come to decide who can qualify as an observer we become entangled in yet another set of questions of a correspondingly confused nature. Who is honest and without bias? Anybody who follows either current science or the past history of science cannot fail to conclude that all have been biased in one way or another (the reviewer includes himself among those who are biased). New ideas, such as those of Einstein, have encountered tremendous bias against them, from almost all the 'reputable' physicists of the time. And then, whose senses are not preconditioned to influence future sensations? It has been established by many observers that it is impossible for our perceptions not to be influenced in a very fundamental way by all that we have learned, especially in the first few years of our lives.¹ Most of this influence is now forgotten, but it operates tacitly, unconsciously and all-pervasively. This leads to what is by far the most dangerous kind of bias—i.e. the bias of a man who believes that he is unbiased.

When the reviewer considers how entangled and confused all these questions are, he doubts that to start with sensations, subject, etc., will really give rise to greater clarity and deeper understanding than is possible with our current starting points. Rather, he gets the impressions of a proposal that a vast symbolic structure of great precision (resembling a modern steel and glass skyscraper) is to be erected on top of a quagmire. Before he feels impelled to 'plunge in' and steep himself in the SL language (as advised by the author) he would therefore like to see some reason to believe that he is not jumping from the 'frying pan' of our current conception of the object, as basic, into the 'fire' of the author's conception of the subject, and his sensations, as basic.

DAVID BOHM
Birkbeck College, London.

¹ A condensed account of some of these observations, along with a detailed set of references, is to be found in D. Bohm: *Special Theory of Relativity* (Benjamin; N.Y. 1965).

Israel Scheffler: *Science and Subjectivity*. (Indianapolis: The Bobbs-Merrill Company, Inc., 1967. \$2.25.) Pp. vi+132.

Recent literature, associated particularly with the work of Feyerabend, Hanson, and Kuhn, has exhibited much dissatisfaction with standard deductive and objectivist views of science, and has argued for the understanding of science in terms of various kinds of relativism: of theory and observation, of meaning and theoretical framework, and of rationality and historical paradigm. Professor Scheffler sees in these writings a dangerous (even *morally* dangerous) surrender to subjectivism, and seeks to re-establish the objective claims of science, while doing full justice to the insights which such forms of subjectivism undoubtedly contain. He adopts a useful dialectical method; bringing into the argument not only recent and fashionable exponents on both sides, but also some illuminating discussions from older writers including C. I. Lewis, Reichenbach, Neurath and Schlick. He is as successful in his aim as it is possible to be in five short chapters, and it is a high compliment to his achievement that he states the opposing views so cogently that one remains dissatisfied with his attempted resolutions.

The most important and troublesome of the cluster of problems dealt with is that of *meaning*. This can indeed be said to subsume the problem of what it is to be an *objective observation*, for as Scheffler argues in his second chapter, observations have to be reported in terms of some *category system*, that is, they presuppose an inter-subjective language. Moreover, observation reports are not infallible and category systems are not unchangeable. They share these instabilities with theories. Thus the relation between theory and observation can be seen to be essentially the same as that between different theoretical systems, and here the problem of meaning arises in sharp paradoxical form.

The paradox of meaning can be put briefly as follows: (1) the meanings of terms in a theoretical system are dependent on the whole theoretical context, and change with change of system (for example 'mass' in Newtonian mechanics means something different from 'mass' in relativistic mechanics); (2) therefore no term P in theory T_1 can be said to mean the same as P in theory T_2 ; (3) therefore no statement in T_1 can entail, be consistent with, contradict, or be evidence for, any statement in T_2 . But this of course is a *reductio ad absurdum* of the whole scientific enterprise, because it implies that no corrections or improvements of theories can take place; there can be only a sequence of theories, perhaps psychologically and historically conditioned, but unrelated by any kind of logical or rational tie.

Scheffler's suggested solution is ingenious, but not I think ultimately acceptable. It is intuitively fairly clear that the apparent paradox has arisen because of an illegitimate shift in the meaning of 'meaning' from (1) to (3), and that the sense of 'meaning' which makes (1) plausible is not the sense required if (3) is to follow. The difficulty is to make this intuition precise. Scheffler suggests that 'meaning' should be construed in the classic Fregean manner as having two separable components: 'sense' and 'reference'. (1) and (2) may be regarded as the assertions that the sense, or definition, or synonymy relations of terms differ in different theories, but in considering the logical relations of deducibility, consistency, contradiction, and so on, it is sameness of reference or extension

that is solely involved. Difference of sense does not imply difference of reference, hence (3) does not follow from (1) and (2). Thus Scheffler claims to reconcile variance of meaning between theories, and between theory and observation, with invariance of reference and hence of logical relations.

Unfortunately this reconciliation does not work. Waiving difficulties about construing sense in terms of definitional synonymy relations (does this imply that the primitives in the definiens have *no* sense?), there is a curious air of instrumentalism about this analysis: '... presumed alterations of sense are of no general scientific interest anyway, deductive systematisations in terms of new theories being prized irrespective of their effects on alleged synonymies of constituent terms' (p. 64). Is the replacement of classical electrodynamics by quantum field theory then of no interest apart from the instrumental reorganisation of observables? More seriously, however, 'same reference' is neither necessary nor sufficient for the logical comparability that is required of different theories. It is not *sufficient*, because the properties ascribed to objects in science are not extensional properties. Suppose two theories T_1 and T_2 are 'about' two quite distinct aspects of a domain of objects: say their colour relations, and their shapes. It may happen that T_1 and T_2 are such that there is an exact one-to-one correspondence between the sets of predicates of T_1 and T_2 respectively, and that, as far as is known, T_1 is true of any set of objects if and only if T_2 is also true of it. Then the corresponding predicates of the two theories have the same referential meaning. But this does not imply that the theories are the same. So long as no predicates are added to their respective predicate-sets, no development of T_1 can be either consistent with nor contradictory to any development of T_2 . In other words, because science is about *intensional* properties, sameness of extension does not suffice for logical comparability. Furthermore, sameness of reference is not *necessary* for logical comparability. Two different theories may make use of different categorisations or classifications of objects: thus Dalton's atoms have different-extensions from Cannizzaro's atoms, yet we want to be able to say of some of Cannizzaro's statements that they entail or contradict some of Dalton's. Scheffler touches upon such a possibility in allowing that different category systems may 'overlap'. But this is hardly more than metaphor. More detailed analysis is required if we are to resolve the subjectivist paradoxes, and it may be conjectured that classic analyses of 'meaning' are not going to help us much.

When all is said, however, this book remains a most admirable and stimulating introduction to several of the most important and controversial issues in current philosophy of science.

MARY HESSE
University of Cambridge

Richard C. Jeffrey: *The Logic of Decision* (New York: McGraw-Hill, 1965, \$7.95). Pp. xi+201.

With the increased recognition of the importance of decision theory for several areas of philosophy as well as for the behavioural sciences there is a need for a

book which sets forth the theory systematically. The central problem of decision theory concerns the situation in which an individual, or a group, must choose one of several alternative courses of action. The decision maker most often makes his choice given only partial information concerning the relevant states of nature and the consequence of each alternative course of action. Decision theory has both its normative aspect and its descriptive aspect. Most treatments are based on the Bayesian principle that people should, in normative decision theory, or do, in descriptive decision theory, perform that act (or one of several) that has maximum expected utility or desirability. A brief general picture of decision theory can be found in Patrick Suppes, 'The Philosophical Relevance of Decision Theory', *Journal of Philosophy*, 63 (1961), 605-14.

R. C. Jeffrey's *The Logic of Decision* provides an elementary introduction to parts of decision theory, develops the theory in an interesting nonstandard way, and discusses its applications to problems concerning induction. In the first three chapters Jeffrey uses the classical approach of F. P. Ramsey and von Neumann and Morgenstern to introduce the basic concepts of decision theory. The classical approach provides a way in which the utilities the decision maker gives to the consequences of the possible courses of action and the probabilities he assigns to the relevant states of nature can be determined from his preference rankings of certain gambles. The bulk of the book, chapters 4 to 10, gives an alternative approach which uses preference rankings of propositions, atomic and those obtained by negation, conjunction, and disjunction, to obtain the utility and probability assignments. In Jeffrey's approach these assignments are made to the same kind of entity, propositions; whereas in the classical approach utilities are assigned to consequences and probabilities to states of nature. Jeffrey is aware of some of the difficulties this gives rise to in ordinary ways of talking; these difficulties are not negligible, but the idea is well worth exploring.

Jeffrey believes that the main advantage his approach has over the classical approach is that it does not make unrealistic requirements on the comparability between consequences. In the classical approach if A and B are two consequences in the decision maker's preference rankings and if there are any gambles on the condition (H) *this coin lands head up* in his preference rankings, then he must also be able to rank the gamble A if H , B if not. However, if A is the consequence *there will be a thermonuclear war next week*, the decision maker is unable to assess the gamble. Jeffrey concludes, 'for the agent to consider that this gamble might be in effect would require him so radically to revise his view of the causes of war and weather as to make nonsense of whatever judgment he might offer'. Jeffrey holds that his own theory, contrary to the classical theory, is noncausal; it does not put demands on the decision maker as illustrated by his inability to assess the above gamble.

In the Ramsey-von Neumann and Morgenstern approach the utilities the decision maker assigns to the consequences of the alternative acts are determined up to a positive linear transformation (that is, two assignments u and U of numbers to represent these utilities are related in such a way that there exists a positive real number a and a real number b such that for every consequence x , $U(x) = au(x) + b$); in this approach the probabilities assigned by the decision maker to the states of nature are completely determined. Jeffrey's formulation yields somewhat weaker results. Two pairs of utility and probability functions

(u, p) and (U, P) satisfying Jeffrey's conditions are related in such a way that there are real numbers a, b, c , and d such that $ad - bc$ is positive and such that for every proposition A with positive probability, $U(A) = \frac{au(A) + b}{cu(A) + d}$ and $P(A) = \frac{p(A)}{cu(A) + d}$. If the preference ranking can only be represented by a utility function that is unbounded both above and below, the classical results are obtained. Unfortunately Jeffrey does not give a set of conditions sufficient for the existence of the utility and probability functions.

In the last two chapters Jeffrey applies his version of decision theory to some questions of philosophical importance. In chapter eleven he considers the case in which an individual changes his degree of belief in one or more propositions on the basis of observation although there might not be any evidential proposition that he learned to be true. The last chapter of the book deals with questions concerning subjective and objective probability and concerning Goodman's 'grue' problem of induction.

I would like to add a few comments in the light of having used the book as a textbook in an elementary course. The first five chapters are a very good introduction to decision theory and to Jeffrey's formulation. However, two or three of the later chapters would have been better left as appendices; for example, the details of the proof of his uniqueness results are interesting to the specialist but are not essential to an 'elucidation of the notions of *subjective probability* and *subjective desirability or utility*'. Some of the exercises provided are excellent, but there are chapters in which the first exercise or two is much more difficult than others that follow. There is no question that this book is useful and provides an interesting alternative to the classical approach to decision theory.

RICHARD E. ROBINSON
University of British Columbia

Robert S. Hartman: *The Structure of Value: Foundations of Scientific Axiology*. (Southern Illinois University Press, 1967.) Pp. xix+384.

This is an imaginative book by a writer who, refreshingly, does not see philosophy as an ivory tower subject. Professor Hartman holds a version of the Russellian thesis that when a topic of discourse finds its feet, it breaks away from philosophy and establishes itself as a science. His obvious example is physical science as it emerged from philosophy in the seventeenth-century revolution. And he holds that our thinking about value at the present time, and in particular about morality, is in the same state as thinking about the natural world before the revolution. Engagingly, he expresses one corollary of his view by writing (p. 47):

Quacks in *natural* philosophy are easily recognised today but were not in the Middle Ages where they occupied respected positions. Quacks in *moral* philosophy are not easily recognised today; they occupy high positions in the academic, political and religious hierarchy.

But, Hartman thinks, the time is ripe for the development of value thinking so that quacks shall be recognisable henceforward.

So far, perhaps, so good. At least it seems that a little air might be let into a subject which needs it. But while Hartman may hold the sympathy of many readers through Part I of his book, few will remain with him as they negotiate the bizarre machinations of Parts II and III. It is not easy to take seriously a 'proof', of the 'Infinite Value of the Human Person' which has as a premiss a theorem of transfinite mathematics (pp. 116-19); nor a 'proof' that aesthetic value is infinitely more valuable than economic value—in the precise sense of "infinitely more" that is defined by the difference between $\aleph_0$ and $\aleph_1$ (p. 114). And claims such as, '*All x ought to be no good . . .*' seems obviously false'; and 'Other propositions seem obviously true, such as *some x ought to be good . . . some x ought to be bad . . .*' (p. 171) display a weird conception of the use of 'ought'.

Hartman is led to think such 'proofs' possible by his mistaken definition of 'good', aided and abetted by his lack of any feeling for plausibility. This mistaken definition, which is the core of his book, enables him to think that he is developing a new subject, value axiology, which stands to value phenomena as mathematics stands to physical phenomena or the theory of harmony to music (p. 30). It is through the development of this new subject that he thinks value discourse can break away from philosophy and establish itself as a science.

The definition of 'good' is given over and over again, each time in a slightly different form, but one of the earlier versions will do for illustration (p. 51):

If there is given an intension containing a number n of predicates (a set of predicates) determining a certain extension (class) then the things belonging to the class (the class members) which have n properties are good such things (good class members); and all those which have less than n predicates are 'less than good' ('fair'), 'not good' ('bad') or the like, such things or class members.

This definition has the consequence that a thing is a good x if and only if it is an x , with the further consequence that there cannot be a fair or bad x . Hartman apparently does not notice this difficulty for a hundred further pages (p. 157) and it is not tackled for forty more. At this point, an untenable distinction is introduced between definitional and expositional properties and we are told that a thing is a good x if and only if it has all the expositional properties of x 's. Examples of expositional properties are given: an expositional property of horses, for example, is 'having horny patches or chestnuts on the inner side of both pairs of legs'. But if one racing man said to another, 'She's a good filly' and the other replied, 'No she's not: no chestnuts', he would not have disputed what the first said. He would be guilty of irrelevance. An elementary reflection such as this shows that Hartman's definition of 'good' will not do; and when this definition is swept away, the whole of Hartman's thesis is swept away with it.

G. R. GRICE
University of East Anglia

Recent Publications

RECENT PUBLICATIONS ON THE PHILOSOPHY OF SCIENCE

ARTICLES

General

- Gerd Buchdahl: Semantic sources of the concept of law. *Synthese* 1 (1967), 54-74
 Peter Caws: Aspects of Hempel's philosophy of science. *Rev. Metaphysics* 20 (1967), 690-710
 Arthur H. Copeland Sr.: Mathematical proof and experimental proof. *Phil. Sci.* 33 (1966), 303-16
 Arthur I. Fine: Consistency, derivability and scientific change. *J. Phil.* 64 (1967), 231-9
 Leon Goldstein: Theory in History. *Phil. Sci.* 34 (1967), 23-40
 Adolf Grünbaum: The status of temporal becoming. *Ann. N.Y. Acad. Sci.* 138 (1967), 374-95
 Herbert Hochberg: Dispositional problems. *Phil. Sci.* 34 (1967), 10-17
 James Hullett and Robert Schwartz: Grue: some remarks. *J. Phil.* 64 (1967), 259-70
 Władysław Krajewski: Four concepts of causality. *Stud. Filozof.* 3 (1966), 77-92
 Irwin C. Lieb: Logic and reality. *Rev. Metaphysics* 20 (1967), 662-89
 D. H. Mellor: Inexactness and explanation. *Phil. Sci.* 33 (1967), 345-59
 D. H. Mellor: Imprecision and explanation. *Phil. Sci.* 34 (1967), 1-9
 Witold Marciszewski: Reduction in the light of an analysis of observation statements. *Stud. Filozof.* 3 (1966), 105-14
 Ernan McMullin: Recent work in philosophy of science. *New Scholast.* 40 (1966), 478-518
 Fred Newman: Two analyses of prediction. *Theoria* 32 (1966), 45-55
 Shozo Onmori: Natural sciences are a natural history. *Ann. Jap. Ass. Phil. Sci.* 3 (1967), 5-18
 K. W. Rankin: Past and future. *Phil. Q.* 17 (1967), 317-33
 G. Schlesinger: What does the denial of absolute space mean? *Aust. J. Phil.* 45 (1967), 44-60
 H. R. Shuford Jr.: Logical constructs. *Scientia* 102 (1967), 38-47
 Herbert A. Simon and Nicholas Rescher: Cause and counterfactual. *Phil. Sci.* 33 (1966), 323-40
 W. A. Suchting: A note on the principle of causality. *Phil. Stud.* 18 (1967), 14-6
 W. A. Suchting: Deductive explanation and prediction revisited. *Phil. Sci.* 34 (1967), 41-52
 Christer Stangenberg: New definitions of scale-types. *Theoria* 32 (1966), 56-61
 Stephen Toulmin: Conceptual revolutions in science. *Synthese* 17 (1967), 75-91

Semantics

- Kazimierz Ajdukiewicz: The problem of empiricism and the conception of meaning. *Stud. Filozof.* 3 (1966), 5-18
 Donald Davidson: Truth and meaning. *Synthese* 17 (1967), 304-23
 Bas C. van Fraassen: Meaning relations among predicates. *Notas* 1 (1967), 161-80
 C. D. Gruender: Science and verification. *Scientia* 150 (1966), 522-7
 Gilbert Harman: Quine on the meaning and existence I. *Rev. Metaphysics* 21 (1967), 124-51
 Jerrold J. Katz: Recent issues in semantic theory. *Fdn. Lang.* 3 (1967), 124-94
 Jerrold J. Katz: Some remarks on Quine on analyticity. *J. Phil.* 64 (1967), 36-51
 Barbara Stanosz: The problem of intentionality. *Stud. Filozof.* 3 (1966), 199-212
 Andrzej Zabłudowski: Testability and meaning. II. *Stud. Filozof.* 48 (1967), 65-90

Ontology

- B. A. Brody: Natural kinds and real essences. *J. Phil.* 64 (1967), 431-45
 Terence Parsons: Existential theories of ontological commitment. *J. Phil.* 64 (1967), 446-50
 R. Sundararajan: Language and ontological commitments. *Darshana Int.* 6 (1966), 11-31
 D. W. Theobald: Observation and reality. *Mind* 76 (1967), 198-207.

Epistemology

- A. J. Ayer: Has Austin refuted the sense-datum theory? *Synthese* 17 (1967), 117-40
 Arthur W. Collins: The epistemological status of the concept of perception. *Phil. Rev.* 76 (1967), 436-59
 Brian Ellis: Physical Monism. *Synthese* 17 (1967), 141-61
 Richard E. Grundy: Some comments on confirmation and selective confirmation. *Phil. Stud.* 18 (1967), 19-23
 Norwood Russell Hanson: An anatomy of discovery. *J. Phil.* 64 (1967), 321-51
 Maurice Mandelbaum: Definitions and coherence in sense-perception. *Notus* 1 (1967), 123-38
 Stanislaw Rainko: Diachronic epistemology: outline of problems. *Stud. Filozof.* 48 (1967), 3-42
 Peter Unger: Experience and factual knowledge. *J. Phil.* 64 (1967), 152-73
 J. Zeman: The gnoseological channel. *Scientia* 3-4 (1967), 141-9

Logic

- J. C. D'Alessio: On subjunctive conditionals. *J. Phil.* 64 (1967), 306-9
 P. T. Geach: Identity. *Rev. Metaphysics* 21 (1967), 3-12
 Jean van Heijenoort: Logic as calculus and logic as language. *Synthese* 17 (1967), 324-30
 A. N. Prior: Stratified metric tense logic. *Theoria* 33 (1967), 28-38
 Daniel S. Robinson: Royce's contributions to logic. *Rev. Int. Phil.* 21 (1967), 60-76
 Ernest Sosa: Hypothetical reasoning. *J. Phil.* 64 (1967), 293-305

Probability

- Johannes Gordesch and Hans Georg Knapp: Zur Logischen Problematik der statistischen Hypothesenprüfung. *Phil. Nat.* 10 (1967), 42-55

Physics and Mathematics

- Peterka Antonin: Forms and kinds of matter and its states. *Acta Polytech.* 5 (1966), 49-60
 Zygmunt Chylinski: Localization problem and space-time continuum. *Stud. Filozof.* 3 (1966), 19-42
 Leon Cohen: Can quantum mechanics be reformulated as a classical probability theory? *Phil. Sci.* 33 (1966), 317-22
 A. G. Drachmann: Archimedes and the science of physics. *Centaurus* 12 (1968), 1-11
 E. M. Fournier d'Albe: The relativity of size. *Scientia* 7-8 (1967), 291-303
 Adolf Grünbaum: The denial of absolute space and the hypothesis of a universal nocturnal expansion: a rejoinder to George Schlesinger. *Aust. J. Phil.* 45 (1967), 61-91
 Armin Hermann: Die frühe Diskussion zwischen Stark und Sommerfeld. *Centaurus* 12 (1968), 38-59
 M. Moser, B. Jahos and H. Schleichert: Gespräch über das Uhrenparadoxon. *Phil. Nat.* 10 (1967), 23-41
 Ryszard Kerner: On the interpretation of the relativity theory. *Stud. Filozof.* 48 (1967), 151-65
 Hilary Putnam: Time and Physical geometry. *J. Phil.* 64 (1967), 240-7
 Hilary Putnam: Mathematics without foundations. *J. Phil.* 64 (1967), 5-22
 C. W. Reithdijk: A rigorous proof of determinism derived from the special theory of relativity. *Phil. Sci.* 33 (1966), 341-5
 Mendel Sachs: On the elementarity of measurement in general relativity: toward a general theory. *Synthese* 17 (1967), 29-53

- Daniel Sedey: Being partial to objects. *J. Phil.* **64** (1967), 175-84
 Lawrence Sklar: The falsifiability of geometric theories. *J. Phil.* **64** (1967), 247-53
 Jan Jerzy Slawianowski: More on determinism in quantum mechanics. *Stud. Filozof.* **48** (1967), 91-120

Biological, Behavioral and Social Sciences

- Ludwig von Bertalanffy: Mind and body re-examined. *J. Humanist. Psychol.* **6** (1966), 113-38
 Brand Blanshard and B. F. Skinner: The problem of consciousness: a debate. *Phil. Phenom. Res.* **27** (1967), 317-37
 Richard Brandt and Jaegwon Kim: The logic of the identity theory. *J. Phil.* **64** (1967), 515-37
 Andre A. Caloutsis: Quelques considérations sur les implications philosophiques de la médecine psychosomatique. *Evolution Psychiat.* **31** (1966), 197-209
 Arthur B. Cody: Can a single action have many different descriptions? *Inquiry* **10** (1967), 164-80
 Frederick J. Crosson: The concept of mind and the concept of consciousness. *J. Existentialism* **6** (1966), 449-57
 Hubert L. Dreyfus: Why computers must have bodies in order to be intelligent. *Rev. Metaphysics* **21** (1967), 13-32
 David P. Gauthier: How decisions are caused. *J. Phil.* **64** (1967), 147-51
 A. James Gregor: Psychoanalytic disposition terms and reduction sentences. *J. Phil.* **63** (1966), 737-44
 Marjorie Grene: Straus's phenomenological psychology. *Rev. Metaphysics* **21** (1967), 94-123
 Rolf Gruner: Understanding in the social sciences and in history. *Inquiry* **10** (1967), 151-63
 John Kekes: Physicalism, the identity theory and the doctrine of emergence. *Phil. Sci.* **33** (1966), 360-76
 Richard Lichtman: Indeterminacy in the social sciences. *Inquiry* **10** (1967), 139-50
 Alex. C. Michalos: Postulates of rational preference. *Phil. Sci.* **34** (1967), 18-22
 Theodore Mischel: Pragmatic aspects of explanation. *Phil. Sci.* **33** (1966), 46-60
 Nicholas Rescher: Values and the explanation of behaviour. *Phil. Q.* **17** (1967), 130-6
 Fred S. Roberts and Patrick Suppes: Some problems in the geometry of visual perception. *Synthese* **17** (1967), 173-201
 D. Stanley Jones: Kybernetics of the nervous system. *Scientia* **3-4** (1967), 150-65
 Robert J. Swartz: Color concepts and dispositions. *Synthese* **17** (1967), 202-22
 Charles Taylor: Mind-body identity: a side-issue? *Phil. Rev.* **76** (1967), 201-13
 Jenny Teichmann: The contingent identity of mind and brains. *Mind* **76** (1967), 404-15
 Ludwig Wittgenstein: Bemerkungen über Frazers 'The Golden Bough'. *Synthese* **17** (1967), 233-53

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

1968 ANNUAL CONFERENCE

The Conference will be held at the University of Exeter from Friday, 20 September 1968 to Sunday, 22 September.

1. *Friday, 20 September, 8.00 p.m.*
Mr P. H. Matthews: The Problem of Verification in Linguistics
Commentator: Mr D. C. Ferris
2. *Saturday, 21 September, 10.00 a.m.*
Miss Margaret Boden: Some Philosophical Implications of Recent Brain Research
Commentator: Professor Robert Ackermann
3. *Saturday, 21 September, 8.00 p.m.*
Mr C. B. Wright: The Structure of Scientific Theories
Commentator: Dr A. J. Lyon
4. *Sunday, 22 September, 10.00 a.m.*
Professor Roger Penrose: A Theory of Space-Time
Commentator: Professor D. Bohm

PROGRAMME OF MEETINGS 1968-69

21 October

Dr D. Knight: Philosophy and the Chemists

18 November

Dr I. Lakatos: The Role of Criticism in Science: Popper *versus* Polanyi and Kuhn

27 January

Mr R. Harré: Qualities and Powers

10 February

Dr J. Giedymin: Some Problems of Confirmation Theory

24 February

Mr J. L. Mackie: Presidential Address

17 March, 4.30 p.m.

ANNUAL GENERAL MEETING

Professor L. Weiskrantz: A Psychologist's View of the Methodology of Behaviour Studies

5 May

Dr H. Post: The Pauli Principle

26 May

Dr R. Young: Persons, Organisms and Primary Qualities

16 June

Professor M. Strauss: Problems of Inter-Theory Reduction

Meetings are held in the Joint Staff Common Room at University College London at 5.30 p.m.

The Correspondence of Isaac Newton

Volume IV 1694-1709

Edited by J. F. SCOTT

This volume contains discussions of Newton's lunar theory, Halley's comets, and the publishing of Flamsteed's *Astronomical Observations*. Newton's draft for the *Optiks* is also included. 'A volume so long awaited, so sorely needed, so excellently produced, and so endlessly fascinating.'

Times Literary Supplement

£11 11s. net

Public Knowledge

The Social Dimension of Science

JOHN M. ZIMAN

A practising scientist and gifted expositor presents a new view of the nature of science and how it works. His essay, which is written in plain English and requires only the slenderest technical knowledge of science, can (and should) be read by any educated person; as he says, 'all genuine scientific procedures of thought and argument are essentially the same as those of everyday life'. *Cloth 22s. 6d. net; paperback 10s. 6d. net*

Assumption and Myth in Physical Theory

H. BONDI

Professor Bondi discusses some of the myths that have grown up around various scientific theories and ideas, particularly special relativity and Mach's principle.

17s. 6d. net

Scientific Principles and Moral Conduct

J. B. CONANT

Dr Conant divides all experience into three categories, those of nature, human nature and religious experiences. He shows that the first is different in kind from the other two, and concludes that ethics cannot be based on science.

6s. net

Freedom of Action in a Mechanistic Universe

D. M. MacKAY

In the twenty-first Eddington Memorial Lecture, Professor MacKay takes a fresh look at the relation between brain mechanism, human responsibility and the old argument about free will. He demonstrates that there is no antithesis between the mechanistic theory of brain processes and freedom of decision.

5s. net

CAMBRIDGE UNIVERSITY PRESS

THE AUSTRALASIAN JOURNAL OF PHILOSOPHY

EDITED BY A. K. STOUT

VOL. 46

MAY, 1968

No. 1

REVIEW ARTICLE: Encyclopedia of Philosophy

ARTICLES

- | | |
|-----------------|---|
| N. FOTION | The Moral Mode of Speech |
| JAAKKO HINTIKKA | Epistemic Logic and the Methods of Philosophical Analysis |
| A. OLDING | The Ontological Status of Secondary Qualities |
| WILLIAM RUDDICK | On Responses and Reactions |
| KAREL LAMBERT | On the No-Type Theory of Significance |

NOTES AND NEWS

Published in May, August and December; Annual Subscription £1 2s. Sterling (£1 to members of the Australasian Association of Philosophy, which you are invited to join). A sample copy and other information may be obtained from the Department of Philosophy, University of Sydney, N.S.W.

Philosophy and Phenomenological Research

AN INTERNATIONAL QUARTERLY

Edited by MARVIN FARBER

JUNE, 1968, VOLUME XXVIII, No. 4

CONTENTS

I. ARTICLES

- The Uses of Inconsistency, by RUBIN GOTESKY
The Pictures in the Head of a Man Born Blind, by CHARLES F. KIELKOPF
Creativity, Correspondence, and Statements About the Future, by DOUGLAS BROWNING
The Logic of Parts and Wholes in Husserl's *Investigations*, by ROBERT SOKOLOWSKI
Hume on Passions, by PAUL DIETL

II. DISCUSSIONS

- The Importance of Intrinsic Value, by ROBERT EDGAR CARTER
Art and the Language of Action, by JACK PUSTILNIK
Pleasure—Disposition or Episode?, by WARREN S. QUINN
A Note on William James and Indian Philosophy, by DALE RIEPE

III. REVIEWS, COMMUNICATIONS, NOTES AND NEWS, AND RECENT PUBLICATIONS

- Subscription prices \$6.00 per year for individuals; \$7.50 per year for institutions.
Separate issues, \$1.65 and \$2.00 respectively.
Postage and handling for back issues only.

A circular listing the main contents of the Journal since 1940 will be sent upon request

Address all communications to

PHILOSOPHY AND PHENOMENOLOGICAL RESEARCH

State University of New York at Buffalo, Buffalo, New York, 14214

MIND

A Quarterly Review of Psychology and Philosophy

edited by Professor GILBERT RYLE

JULY 1968

- I. The Seventh Letter of Plato
II. Proper Names
III. On the Conception of Polylogism
IV. The Precepts of Justice
V. Unfulfilled Conditionals and the Truth of their Constituents
VI. Value Judgments and Action
VII. Discussion Notes:
 Hare's Debtors
 The Regress Arguments in the *Republic*
 Chisholm on Knowledge and Principles
 Actions, Brain-Processes, and Determinism
 On a Finitist "Solution" to some Zenonian Paradoxes
 R. M. Hare and Moral Reasoning
 Zeno and Continuity
 Has Lee Finally caught Zeno?
VIII. Critical Notices:
 The Commonplace Book of G. E. Moore (1919-1953) Edited by Casimir Lewy
 Lectures on Philosophy. By G. E. Moore. Edited by Casimir Lewy (G. J. WARNOCK)
 Contemporary Moral Philosophy. By G. J. Warnock (R. M. HARE)
IX. New Books
X. Notes
- M. LEVISON, A. Q. MORTON AND A. D. WINSPEAR
RICHARD CAMPBELL
GEORGE BRUTIAN
J. I. MACADAM
JONATHAN HARRISON
L. W. SUMNER
D. P. GAUTHIER
D. R. DUFF-FORBES
BERNARD ROSEN
R. E. EWING
G. A. WEDEKING
G. H. HARMAN
LEIGH VAN VALEN
DONALD MOOR

Annual Subscription 25s, post free; single copy 7s 6d, plus postage

Published for MIND ASSOCIATION by

BASIL BLACKWELL

49 BROAD STREET OXFORD

THE BRITISH JOURNAL FOR THE HISTORY OF SCIENCE

CONTENTS OF VOL. IV, PART I, No. 13

JUNE 1968

Main Articles:

- The Background to the Discovery of Dulong and Petit's law
The cause of refraction in Medieval Optics
The third law in Newton's mechanics

ROBERT FOX
DAVID C. LINDBERG
RODERICK W. HOME

Notes and Communications:

Lavoisier's early career in science. An examination of some new evidence.

Babbage and Moll on the state of science in Great Britain. A Note on a Document

Chemistry in space and the complex atom

J. B. GOUGH
NATHAN REINGOLD
W. V. FARRAR

Book Reviews:

PUBLISHED BY THE BRITISH SOCIETY FOR THE
HISTORY OF SCIENCE

Registered Office;

SCIENCE MUSEUM, SOUTH KENSINGTON, LONDON, S.W.7

INQUIRY

An Interdisciplinary Journal of
Philosophy and the Social Sciences

Edited by ARNE NAESS

Volume 11

Number 2

Summer 1968

Duties to Oneself and the Concept of Morality

Paul D. Eisenberg

Historicism and Historical Laws of Development

Laird Addis

Do Anthropologists Become Moral Relativists By Mistake?

V. A. Howard

Scepticism and Conditions for Description

Peter Zinkernagel

Sociology without Sociology

Hans J. Hammell and Karl-Dieter Opp

Remarks on Empirical Semantics

Jan Berg

Discussion:

Scientific Ethics and the Community

R. J. Brownhill

Review discussion:

Jan Narveson: *Morality and Utility*

Timothy L. S. Sprigge

*Published quarterly by Universitetsforlaget, P.O. Box 307,
Blindern, Oslo 3, Norway*

Annual Subscription: 35s., \$6.00, postage included

THE REVIEW OF METAPHYSICS

A PHILOSOPHICAL QUARTERLY

Edited by RICHARD J. BERNSTEIN

Volume XXI, No. 4

Issue No. 84

June 1968

ARTICLES

The Self as Source of Meaning in Metaphysics, W. NORRIS CLARKE, S.J.

Dare to be Wise, RICHARD TAYLOR

The One and the Many, GARETH B. MATTHEWS AND S. MARC COHEN

CRITICAL STUDIES

Greek Historians, LEO STRAUSS

Philosophical Analysis and Historical Understanding, LOUIS O. MINK

Paul Weiss's Philosophical Journal, ANDREW J. RECK

DISCUSSION

Intelligence, Bodies, and Digital Computers, KENNETH M. SAYRE

EXPLORATION

The Place of Les Mots in Sartre's Philosophy, JOSE HUERTA-JOURDA

BOOKS RECEIVED

Summaries and Comments, HOWARD P. KENIG AND STAFF

PHILOSOPHICAL ABSTRACTS

ANNOUNCEMENTS

Annual Subscription \$6.00

Student Subscription \$4.00

LYMAN-BEECHER HALL, HAVERFORD COLLEGE, HAVERFORD, Pa. 19041 USA

Readers resident in Great Britain may subscribe at a special rate through our British representative, Professor Stephan Körner, Department of Philosophy, The University, Bristol 8 England

“SCIENTIA”

INTERNATIONAL REVIEW OF SCIENTIFIC SYNTHESIS

(1968 - 62nd Year)

Scientific Committee

G. ABETTI - G. B. BONINO - P. CALDIROLA - L. CALIFANO - R. CANESTRARI -
G. COLONNETTI - B. DEFINETTI - G. MERLA - G. MONTALENTI - E. PADOA - G. SANSONE
A. SESTINI - V. SOMENZI - S. TONZIG.

Editor N. BONETTI

“SCIENTIA” publishes original articles by leading workers in all the fields of science and philosophy, from all countries.

It is a magazine of synthesis, that is, of the convergence of the various branches of knowledge.

It is a magazine of multidisciplinary information; it promotes the “bridge” disciplines, which are the most active in the progress of research.

It encourages the building of a culture fully adherent to modern scientific and philosophic standards.

Published every two months, each issue containing approximately 200 pages.

Articles are published in the original language of the author i.e. English, Italian, French, Spanish and German. Every issue contains a supplement with a French translation of the non-French articles. Each issue contains: Book reviews - Analysis - Chronicle.

SUBSCRIPTION RATES: Italy Lire 15.000 - Abroad \$32.00 or £13
Current issue: Italy Lire 3.000 - Abroad \$6.50 or £2.

Back issues: Italy Lire 500 - Abroad \$1.50 or 11s. (in stamps).

“Scientia” representative in U.K.: Atlas Publishing & Distributing Co. Ltd. -
334 Brixton Road - London S.W.9

“SCIENTIA” - Via Donizetti, 55 - 20122 Milano (Italy)

JOURNAL OF THE HISTORY OF IDEAS

A Quarterly devoted to Intellectual History

VOLUME XXIX

APRIL-JUNE 1968

NUMBER 2

TABLE OF CONTENTS

ARTICLES

Pomponazzi and 'Double Truth'	Martin Pine
Ward's 'Pill and Drop'	Marjorie H. Nicolson
Bentham and Penal Reform	Barbee-Sue Rodman
Bourgeois and Proletarians	Gerald A. Cohen
Marxism, Nationalism, and Russia	Neil A. Martin
The Frontier Thesis and the Depression	Steven Kesselman

NOTES

Caxton's Book on 'Chesse'	David Antin
The Beast-Machine in English Literature	Wallace Slugg
New Light on an Old View (Combe on Phrenology)	A. Cameron Grant

REVIEW

<i>The Complete Works of St. Thomas More</i> , Volume 4, <i>Utopia</i> , edited by Edward Surtz and J. H. Hexter	Reviewed by Arthur B. Ferguson
---	--------------------------------

Subscription Rates (outside U.S.A. post paid): \$7.50 for one year;

\$2.00 for single issue

Payable to

JOURNAL OF THE HISTORY OF IDEAS, INC.

The City College, New York, N.Y. 10031

HISTORY OF SCIENCE

An Annual Review of Literature, Research and Teaching

Edited by A. C. CROMBIE and M. A. HOSKIN

CONTENTS OF VOLUME 6, 1967-68

ARTICLES:

Rosalind C. H. Tanner—The Study of Thomas Harriot's Manuscripts: I. Harriot's Will
Jon V. Pepper—The Study of Thomas Harriot's Manuscripts: II. Harriot's Unpublished Papers

A. P. Youschkevitch—Recherches sur l'histoire des mathématiques au Moyen Age dans les pays d'Orient

D. T. Whiteside—A Face-lift for Newton: Current Facsimile Reprints

J. A. Lohne—The Increasing Corruption of Newton's Diagrams

John R. Levene—Sources of Confusion in Descartes's Illustrations

ESSAY REVIEWS:

Carl B. Boyer—The Making of a Mathematician

C. Webster—The Origins of the Royal Society

A. G. Debuss—Alchemy and the Historian of Science

J. S. Wilkie—Preformation and Epigenesis

Richard S. Westfall—The Science of Optics in the Seventeenth Century

W. H. Brock—The First Principles of the History of Chemistry

Erwin H. Ackerknecht—The Total Harvey

Published by

W. HEFFER & SONS LTD., Cambridge, England

SYNTHESE

An International Journal for Epistemology, Methodology and Philosophy of Science

Editor: JAAKKO HINTIKKA

University of Helsinki and Stanford University

Board of Consulting Editors: David M. Armstrong, Yehoshua Bar-Hillel, Arthur W. Burks, Robert S. Cohen, Donald Davidson, Dagfinn Føllesdal, Hans Freudenthal, Béla Juhos, Jerrold J. Katz, Imre Lakatos, G. Nuchelmans, Wesley C. Salmon, Wolfgang Stegmüller, Erik Stenius, Patrick Suppes, Klemens Szanawski, H. Törnebohm, Marx Wartofsky.

Contents of Volume 18, No. 1, January 1968

The Structure of Logical Probabilities by *Jens Erik Fenstad*

The Whole Truth by *Richard C. Jeffrey*

On the Confirmation of Hypotheses about Regions of Existence by *Håkan Törnebohm*

Causal Connection by *William Ruddick*

A Generalization of the Principle of Causality, Which Makes It Applicable to Evolutionary Systems by *Claude Hillinger*

Causality Requirements and the Theory of Relativity by *Peter Haves*

No more Discovery in Physics? Review Essay by *Joseph Agassi*

Reviews

Books Received

Annual Subscription Price Dfl. 60.—(\$17.00)

Personal Subscription Price Dfl. 35.—(\$9.80)

D. REIDEL PUBLISHING COMPANY
DORDRECHT-HOLLAND

Revue Internationale de Philosophie

Editorial Board: Profs. Ph. DEVAUX, H. GOUHIER, Ch. PERELMAN and Mrs. Simone SERVAIS

Secretary: 99 avenue de l'Université, Bruxelles 5

Each number is devoted to a particular movement, a particular philosopher, or a particular problem. As complete a bibliography as possible is given.

Truly international in spirit, the *Review* chooses its collaborators among scholars of all countries. The articles are written in English, French, German, Italian, or Spanish.

Publishes 4 numbers annually. (Europe: 350 Belgian francs; other countries: 400 Belgian francs.)

Coming numbers for 1968

Philosophie du Langage au XVIII^e Siècle, Eug. Dupréel, L'Humanisme
L'Analogie

FOUNDATIONS OF LANGUAGE

International Journal of Language and Philosophy

Editorial Board: Morris Halle (MIT), Peter Hartmann (Münster/W), K. Kunjuni Raja (Madras), Benson Mates (Univ. of California), J. F. Staal (Amsterdam), Pieter A. Verburg (Groningen), John W. M. Verhaar (Nijmegen)

VOLUME IV, No. 1, JANUARY 1968

Articles: Wolfgang Iser: Wolfgang Steinitz zum Gedenken — George Lakoff: Instrumental Adverbs and the Concept of Deep Structures — Paul Kiparsky: Tense and Mood in Indo-European Syntax — Ian L. Yen: Two Measures of Economy in Phonological Description — *Reviews* by: J. E. B. Gray, G. J. M. Bartelink, C. A. van Peursen, F. R. Palmer, Robert A. Hall, Jr., F. van der Blij, Henry M. Hoenigswald and Martin Hollis — *Publications Received*

FOUNDATIONS OF LANGUAGE is published quarterly

Annual subscription price /42.—(\$11.75)

D. REIDEL PUBLISHING COMPANY / DORDRECHT - HOLLAND

METAPHYSICS AND MEASUREMENT

ESSAYS IN THE SCIENTIFIC REVOLUTION

ALEXANDRE KOYRE

These six essays centre on Professor Koyre's great theme: the relative importance of metaphysics and observation, with controlled experiment a kind of marriage of the two. *Metaphysics and Measurement* contains the following essays: Galileo and Plato; Galileo and the Scientific Revolution; the *De Motu Graviorum* of Galileo; An Experiment in Measurement; Pascal Savant; Gassendi and Science In His Time.

42s

CHAPMAN & HALL

dialectica

Revue trimestrielle de philosophie de la connaissance

d
i
a
l
e
c
t
i
c
a

- s'attache à promouvoir les principes directeurs de la *méthodologie ouverte*. Image vivante du développement de la philosophie ouverte, elle publie sur cette dernière et ses thèses essentielles des éditoriaux, des articles de fond, les comptes rendus des débats et des discussions sur ses tendances et ses engagements actuels
- organise des débats et publie des enquêtes sur les questions essentielles de la connaissance moderne (voir par exemple les numéros sur l'idée de complémentarité avec la participation de cinq prix Nobel de physique, sur les fondements de la psychologie, du calcul des probabilités, sur la notion d'analogie, le fait et le droit, l'épistémologie de la physique quantique)
- se prête à la remise en cause des bases des disciplines abstraites et le fondement des disciplines expérimentales
- commente les publications récentes et accorde l'hospitalité de ses colonnes aux discussions de principe à leur sujet
- appelle à la collaboration tous ceux qui entendent participer à la synthèse et à l'intégration de l'ensemble des connaissances modernes
- procède actuellement à une enquête sur la question brûlante de la rénovation de l'enseignement des mathématiques, de la physique et de la biologie

dialectica en un mot est l'organe vivant et la libre tribune de la *méthodologie* et de la *philosophie ouvertes*

Comité directeur: FERDINAND GONSETH, Lausanne
PAUL BERNAYS, Zurich
H. KÖNIG, Wabern/Berne
P. E. FULET, Lausanne

Rédaction: 12, chemin du Muveran, 1000 Lausanne (Suisse) Vente et abonnements: W. Roesch & Cie, Monbijoustrasse 9, 3000 Berne (Suisse)

Prix: Suisse Fr 9.- le numéro Etranger: Fr 10.- le numéro
Abonnement: Suisse Fr 30.- Etranger: Fr 35.-

THE BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

Committee 1968-69

President: Mr J. L. Mackie

Vice-President: Dr G. J. Whitrow

Hon. Secretary: Mr R. Harré

Hon. Treasurer: Dr W. Mays

Hon. Secretary of the Northern Group: Dr N. Wetherick

Ordinary Members: Dr J. Giedymin, Mr A. E. Musgrave, Mr H. D. Sluga

EDITORIAL BOARD OF THE JOURNAL

Professor J. D. Bernal

Professor R. B. Braithwaite

Professor H. Dingle

Professor Sir John Eccles

Professor Sir Harold Jeffreys

Professor W. C. Kneale

Professor S. Körner

Professor M. Polanyi

Professor Sir Karl Popper

Earl Russell

Dr G. J. Whitrow

Mr L. L. Whyte

Professor J. H. Woodger

The British Journal for the Philosophy of Science is the official organ of the Society, whose purpose is described by the following statement from its Constitution and Rules :

'The purpose of the Society is to study the logic, the method and the philosophy of science as well as those of the various special sciences, including the social sciences.'

The Society holds meetings, at which papers are read and discussed, once a month during the academic year. These are held, by kind permission, in University College, London.

Ordinary membership is open to all persons who have been proposed by two members, approved by the Committee, and elected at an ordinary meeting. The annual subscription is 20s. Those who wish to join should write to the Hon. Secretary at Linacre College, Oxford.

The Northern Group of the Society meets regularly at various centres in the north of England. Those who wish to join it should write to the Hon. Secretary of the Northern Group at the Undergraduate School of Studies in the Social Sciences, University of Bradford, Bradford 7.

Brit. J. Phil. Sci. Vol. 19 No. 2 August 1968

The British Journal for the Philosophy of Science

15 NOV 1968

ARTICLES

- 93-108 JAMES LEACH Explanation and Value Neutrality
109-122 KEITH GUNDERSON Robots, Consciousness, and Programmed Behaviour
123-143 I. J. GOOD Corroboration, Explanation, Evolving Probability, Simplicity and a Sharpened Razor

DISCUSSIONS

- 145-152 DAVID MILLER The Straight and Narrow Rule of Induction: A Reply to Dr. Bub and Mr. Radner
152-155 GEORGE BERGER The Conceptual Possibility of Time Travel
155-156 J. R. LUCAS Human and Machine Logic: A Rejoinder
156-157 I. J. GOOD The White Shoe *qua* Herring is Pink

159-180 REVIEWS

181-183 RECENT PUBLICATIONS

184 ANNOUNCEMENT



CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, N.W.1.

American Branch: 32 East 57th Street, New York, N.Y. 10022

Single issues 20s. net in U.K. (\$2.75 in U.S.A.)

Subscription price 60s. net in U.K. (\$9.50 in U.S.A.)

Volume 19 Number 3

November 1961



The British Journal for the Philosophy of Science



CAMBRIDGE UNIVERSITY PRESS

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

Editors: MARY B. HESSE, D. H. MELLOR

Review Editors: B. M. FLOOD, G. M. K. HUNT

Subscriptions

The *Journal* is published quarterly (May, August, November and February). The subscription price is 60s. net in U.K. (U.S.A. \$9.50) per volume post free; single issues 20s. net in U.K. (U.S.A. \$2.75) plus postage. The subscription is 40s. (U.S.A. \$7.00) for members of the British Society for the Philosophy of Science, who should order direct from the publisher. Orders may be sent to any bookseller or direct to Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1. Orders from U.S.A. and Canada should be sent to Cambridge University Press American Branch, 32 East 57th Street, New York, N.Y. 10022, U.S.A. Second class postage paid at New York, N.Y.

Notes to Contributors

Articles and correspondence should be sent to the Editors, and books for review to the Review Editors, at the Whipple Museum, Free School Lane, Cambridge. Articles should not normally exceed 5,000 words, and should be accompanied by abstracts of about 100 words which will then appear in *Dialectica*, *Synthese* and *Philosophy of Science*.

All contributions should be typed in duplicate on one side only with wide margins and double-spaced. Footnotes, which should be as few as possible, should be typed separately, double-spaced at the end of the text. They should be referred to by consecutive numbering throughout the text, in which they will be inserted in page proof. Bibliographical references should be listed alphabetically at the end, and reference made to this list in the text rather than in footnotes. The style of references should in future follow the Harvard system.

Long quotations are printed in small type without quotes, and should be so marked in the typescript. Diagrams, tables, and illustrations should be on separate sheets, with their desired position in the text clearly indicated. Symbolic formulae appearing as separate lines should be clearly distinguished in the text, and foreign letters (Greek, German, etc.) indicated in the margin. On matters of style and layout, authors should consult CHAUNDY, T. W. *et al.* (1964). *The Printing of Mathematics*. London: Oxford University Press.

Offprints of articles should be ordered from the Editor when corrected galley proofs are returned. 50 are supplied free; the price of extra copies is obtainable from the publisher.

Advertising

Particulars of advertising in the *Journal* may be obtained from Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1.; or in U.S.A. and Canada, Cambridge University Press American Branch, 32 East 57th Street, New York, N.Y. 10022, U.S.A.

Hidden Variables and the Copenhagen Interpretation—A Reconciliation¹

by JEFFREY BUB



1 INTRODUCTION

One of the most interesting features of the quantum theory from the philosopher's point of view is the so-called *measurement problem*.² There are various ways of formulating this problem, either with reference to the time evolution of an individual system, or in terms of the statistical properties of ensembles. Basically, the problem is to provide an explanation for the fact that within the framework of the quantum theory physical processes involving measurement differ fundamentally from physical processes which do not involve measurement. This is in sharp contrast to the classical theory, according to which a measurement process is not different, in principle, from any other physical process. 'Providing an explanation' in this context means either 'providing an interpretation of the quantum theory', or 'providing a new theory with its own interpretation'.

The Copenhagen interpretation is generally regarded as the orthodox interpretation of the quantum theory. As expressed by Bohr, this view emphasises a 'feature of wholeness typical of proper quantum phenomena'³—the integral relationship between the measuring instrument and the system measured. The 'feature of wholeness' is introduced via an argument which involves the assumption that unambiguously communicable information about what is observed is necessarily described in terms of the concepts of classical physics, or refinements of classical concepts. This argument leads to the conclusion that the quantum theoretical description

Received 1 September 1967

¹ This article is largely a re-statement and an elaboration of certain ideas implicit in a previous article on hidden variables and the measurement problem, of which Professor D. Bohm, of Birkbeck College, London University, was the principal author. A further correspondence with Professor Bohm has contributed considerably to clarifying this point of view. (Needless to say, Professor Bohm bears no responsibility for any inadequacies peculiar to this presentation.) Acknowledgements are also due to Professor C. Alden Mead, of the Department of Chemistry, University of Minnesota, and to Mr William Demopoulos, of the Minnesota Center for the Philosophy of Science, who provided many helpful conversations and suggestions.

² D. Bohm and J. Bub, *Rev. Mod. Phys.* 38 (1966), 453. This article lists some references.

³ N. Bohr, 'Quantum Physics and Philosophy—Causality and Complementarity' first essay in *Essays 1958-1962 on Atomic Physics and Human Knowledge* (London: Interscience, 1963), p. 4. Apparently Niels Bohr regarded this essay as an especially clear and concise expression of his own views on the interpretation of the quantum theory (see the Preface by Aage Bohr). All passages quoted from Bohr are exclusively from this source.

is statistical in an irreducible sense, i.e. that the quantum statistics cannot be related to distributions of variables which have not so far been taken into account. It is therefore generally assumed that the 'hidden variable' approach is an attempt to avoid the 'feature of wholeness', which is regarded as an essential characteristic of quantum phenomena. Hidden variable theories are dismissed as reactionary and fruitless attempts to return to a Newtonian-type determinism.

This attitude towards the hidden variable approach is completely mistaken. It will be argued below that the idea of an *irreducible* quantum statistics follows only from Bohr's explanation of 'wholeness', which involves the questionable assumption of a unique relationship between the objective observational situation and the concepts of classical physics. Even if this assumption is rejected, the 'feature of wholeness' can be retained in a new sense which is quite compatible with the hypothesis of hidden variables. Indeed, as will be made clear below, the deep intention behind the development of hidden variable theories is the realisation of a 'natural philosophy' which incorporates a concept of 'wholeness' as a new ontological thesis.

From this point of view, Bohr's assumption about the significance of classical concepts may be regarded as the conservative aspect of his interpretation of the quantum theory, and his idea of 'wholeness' as the essence of this interpretation. The hidden variable approach is an attempt to extend Bohr's basic insight into the revolutionary significance of the quantum theory, and not a conflicting proposal to restore classical determinism. What is rejected is Bohr's 'conservative thesis'; what is retained, in a new sense, is Bohr's 'revolutionary thesis'.

A hidden variable theory which explicitly incorporates a concept of 'wholeness' in a particularly simple and explicit way has recently been developed,¹ and will be discussed below. The theory has testable consequences which contradict the predictions of the orthodox quantum theory. Specifically, it follows from the theory that an ensemble of ordered sequences of incompatible measurement processes can in principle yield a distribution of results which contradicts the quantum statistics, if the measurement processes follow each other sufficiently rapidly. (See section 5.) This hidden variable solution to the measurement problem provides both a new theory and a new interpretation.

2 THE COPENHAGEN INTERPRETATION

The peculiar status of measurement in the quantum theory is often associated with Heisenberg's uncertainty or indeterminacy relation: $\Delta q \Delta p \geq \hbar$.

¹ D. Bohm and J. Bub, *op. cit.*

This may be interpreted as a reciprocal relationship between the scatter in a distribution of co-ordinate measurements and the scatter in a distribution of momentum measurements on a statistical ensemble of systems all in the same quantum state (h is Planck's quantum of action). A similar relationship holds for energy-time measurements. Such relationships are puzzling if the concept of state in the quantum theory is regarded as analogous to the classical mechanical concept of state, i.e. if all variable quantities associated with a quantum system are simply functions of its quantum state. It will be worthwhile to make this point quite clear with a brief review of the structure of classical mechanics.

For our purposes here, we may regard classical mechanics as characterised by the following features:

(C. 1) The state of a mechanical system is represented by the values of a certain set of basic parameters $\{q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n\}$, the generalised co-ordinate and momentum variables of the system. All other variable quantities associated with the system are functions of these state variables.

(C. 2) A dynamical rule defines the change in the state variables for any specified system and any specified initial values of the state variables.

A physical theory which is based on classical mechanics is both *mechanistic* and *deterministic*. It is mechanistic in the sense that all change in a specified physical system is referred to the change in a certain set of basic variables, so that all physical properties of the system are explained as functions of these variables. This means that the behaviour of any complex system represented by the variables $\{q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n\}$ can be understood as the interaction of (a maximum of n) sub-systems. Each of these component sub-systems preserves its identity as a separate mechanical system, i.e. (C. 1) is applicable both to the composite system as a whole and to its sub-systems as components. (Naturally, in this case the behaviour of each sub-system can not be defined by a dynamical rule referring only to its own state variables, because such systems are not isolated. But the change of state of any component sub-system is still a well-defined concept and can be deduced from the dynamical rule for the composite system.) The theory is deterministic because the dynamical rule is not statistical, but relates a definite final state of a specified system to every initial state.

Measurement, in principle, is an interaction between two mechanical systems, S_1 and S_2 , such that the dynamical rule for the composite system ($S_1 + S_2$) is of such a form that it is possible to relate changes in the variables associated with S_1 to changes in the variables associated with S_2 . (See appendix I.) The interaction not only makes no difference to the simultaneous and

precise measurability of the position and momentum variables, *it is a necessary prerequisite for the measurement.* (Of course, this conception of measurement refers only to the relationship between two interacting mechanical systems and therefore does not answer the question of how values are associated with the co-ordinate and momentum variables of S_1 'by observation' in the first place. But this question goes beyond classical mechanics, which is a theory of the relationships between well-defined mechanical systems.)

The point made above is that the variables q and p (and all other variables) are always well-defined and precisely measurable, in principle, for a classical mechanical system. A distribution of measurement results is always related to a distribution of states in an ensemble of systems. The indeterminacy relations cannot, therefore, be understood as referring to related distributions in the measured values of variable quantities associated with an ensemble of mechanical systems all in the *same mechanical state*. Neither can the quantum state be interpreted as a purely statistical mechanical concept, i.e. as referring to a probability distribution of mechanical states. Although it is possible to map each quantum state on to a distribution function satisfying the indeterminacy relations for co-ordinates and momenta, this function is not necessarily positive and so cannot be interpreted as a probability distribution in general.¹

Heisenberg originally explained these relations as expressing an inherent limitation in the *simultaneous measurability or observability* of the variables q and p (or E and t) for a micro-system, because of the irreducible and uncontrollable disturbance of the system measured by the measuring instrument.² It is usually argued that the irreducibility of the disturbance follows from the necessarily finite energy transfer ($E = h\nu$) in any quantum interaction, but this alone would not provide any explanation for the indeterminacy relations. As the above analysis of the classical mechanical measurement interaction demonstrates, some 'disturbance' of the system measured is necessary for a measurement. The sense in which the disturbance cannot be controlled (and hence cannot be calculated or predicted) is not quite clear from Heisenberg's original arguments based on the γ -ray microscope thought experiment. In fact, this conclusion could only follow from the quantum theory itself, and what is at issue here is just the interpretation of this theory.

Heisenberg's argument is confused because he tacitly asserts the classical

¹ This was first demonstrated by E. P. Wigner in *Phys. Rev.* **40** (1932), 749. See J. E. Moyal, *Proc. Camb. Phil. Soc.* **45** (1949), 99, for a discussion of 'Quantum Mechanics as a Statistical Theory'.

² W. Heisenberg, 'Physical Principles of the Quantum Theory' (Chicago: University of Chicago Press, 1930).

ontology in an argument that explicitly purports to deny it. He treats the measuring instrument and the system measured as two separate mechanical systems in interaction, and he uses the quantum theory to analyse the interaction process. But the quantum theory with its indeterminacy relations is incompatible with the conception of measurement as a mechanical interaction process, symbolised by the diagram of a γ -ray microscope interacting mechanically with an electron particle. Rather than providing an intuitive understanding of the quantum relationships, this sort of argument only serves to obscure the essential difference between the quantum theory and classical mechanics.

Bohr's point of view was much more subtle than that of Heisenberg. Two themes form the basis of his interpretation. Firstly, Bohr emphasised what he referred to as a 'feature of wholeness'¹ in quantum phenomena. Secondly, he insisted on the use of 'plain language, suitably refined by the usual physical terminology',² i.e. the concepts of classical physics, for the description of unambiguous observational information or objective knowledge—information which can be communicated without explicit reference to any individual observer. He argued that 'all unambiguous information concerning atomic objects is derived from the permanent marks—such as a spot on a photographic plate, caused by the impact of an electron—left on the bodies which define the experimental conditions'.³ For Bohr, the 'experimental conditions' are defined by the *classically describable structure* of the macroscopic bodies which constitute the measuring instrument, i.e. by their relative positions and momenta. A physical change in these bodies (e.g. the appearance of a spot on a photographic plate) is related to the *classically describable function* of the measuring instrument (e.g. the appearance of a spot on a photographic plate is interpreted to mean that the instrument has functioned so as to register the position of a particle). Thus, the second theme restricts the structure and function of the measuring instrument at the macro-level to what can be described classically.

Now, the requirement of a *classical description* of the structure and function of the measuring instrument actually provides an explanation for the 'feature of wholeness' in the following way. In the case of phenomena which are inexplicable by the *classical theory*, this requirement implies that the measurement process cannot be understood as a straightforward interaction between two mechanical systems. This is because the system measured is certainly not a classical mechanical system (since the classical theory fails to provide an explanation for the observed behaviour of the

¹ N. Bohr, op. cit. pp. 2 and 4.

² Ibid. p. 3.

³ Ibid. p. 3.

measuring instrument), whereas the structure and function of the measuring instrument is, according to Bohr, necessarily described in purely classical terms. In this sense, the classically describable structure and function of the experimental arrangement is involved in the very *definition* of the 'non-classical objects' investigated, since these 'objects' cannot be conceptually separated in an unambiguous way from the context of a classically specified environment. This, presumably, is what Bohr meant by the 'fundamental distinction between the measuring apparatus and the objects under investigation': *the 'objects' measured in a quantum measurement process have a different status in the theory to the measuring instruments*, i.e. the process is not an interaction between two mechanical systems (as in classical mechanics).

Thus, the 'feature of wholeness' is the *fundamental inseparability or indivisibility* of the relationship between the measuring instrument and the system measured, as a consequence of their crucial difference in the sense that *they belong to different levels of functioning which are not related mechanically*. And this difference is 'a direct consequence of the necessity of accounting for the *functions* of the measuring instruments in purely classical terms, excluding in principle any regard to the quantum of action'.² (Italics inserted.) For Bohr, therefore, a 'quantum phenomenon' is the functioning of a classically describable measuring instrument which cannot be explained according to the classical theory. Consequently, Bohr could confidently assert: 'The very fact that repetition of the same experiment, defined on the lines described, in general yields different recordings pertaining to the object, immediately implies that a comprehensive account of experience in this field must be expressed by statistical laws.'³ In other words, since the observed statistical fluctuations at the macroscopic level are of such a nature that they cannot be explained by the (classical) laws of this level, and since the fluctuation phenomena are necessarily described in terms of the (classical) concepts appropriate to this level, there can be no more complete explanation of the experimental results other than the statistical account in terms of the classical concepts. This explanation is provided by the algorithm of the quantum theory, which is a formal means of relating the statistics associated with different classically describable experimental arrangements.

Although Bohr explains the distinction between measuring instrument and measured system by appealing to the necessity for describing the measuring instrument at the macro-level in purely *classical* terms, the 'feature of wholeness' involved in this distinction may be regarded as the

¹ N. Bohr, p. 3

² Ibid. p. 4.

³ Ibid. pp. 3, 4.

deep 'revolutionary thesis' in Bohr's interpretation, and the assumption of a *classically* describable macro-level as a relatively superficial 'conservative thesis'. The motivation for the assumption of a classically describable macro-level seems to have been a feeling on the part of Bohr that it was necessary to salvage the true 'factual core' of classical mechanics, and to build on it. From the point of view of the hidden variable approach, this attitude represents a methodological or tactical error. (See section 5.) The hidden variable approach develops Bohr's notion that the quantum fluctuations reveal the functioning of a micro-level which is not related to the functioning of the measuring instruments at the macro-level in a simple mechanical way, but drops the restriction to a *classically* describable macro-level. Of course, by dropping part of Bohr's interpretation, the hidden variable approach leads to a new conception of 'wholeness', and so the resulting theory is different from the quantum theory. The primary difference is that the relationship between the micro-level and the macro-level need not necessarily be the statistical relation specified by the quantum theory, which is a consequence of ignoring those features of the macro-level which do not correspond to any classical concept. This does not, however, involve a commitment to a simple mechanical or deterministic relationship.

To sum up: Bohr's interpretation is that there are no 'micro-systems' which can be conceptually isolated from a 'macro-environment' ('revolutionary thesis'), of which the structure and function are classically describable ('conservative thesis'). In other words, there are no 'quantum objects' which constitute the building-blocks of the classical macro-world. There is a 'quantum algorithm' which is applicable to 'quantum phenomena', but it would make no sense to attempt to recover classical mechanics as a special case of the quantum theory, or as a first approximation to the quantum theory (in the sense, for example, that Newton's theory can be recovered from the theory of relativity).

On this view, Heisenberg's indeterminacy relations are 'not concerned with a restriction as to the accuracy of measurements, but with a limitation of the well-defined application of space-time concepts and dynamical conservation laws, *entailed by the necessary distinction between the measuring instruments and atomic objects*'.¹ (Italics inserted.) The functioning of an experimental arrangement so as to record position in a quantum measurement process excludes the simultaneous functioning of the experimental arrangement so as to record momentum (and similarly for energy-time measurements). In the case of an experimental arrangement which can be described as functioning so as to restrict position to a certain range Δq ,

¹ N. Bohr, p. 5.

this arrangement can simultaneously be described as functioning so as to restrict momentum to the range Δp . Statements about space-time measurements are said to be in a *complementary relationship* to statements about momentum-energy measurements. Thus, Heisenberg's relations are interpreted as a quantitative expression of this complementary relationship.

Schumacher has suggested that Bohr intends the 'principle of complementarity' to be regarded as a *definition* of the *quantum description*. Schumacher states the principle as follows: 'To mutually exclusive experimental conditions there corresponds distinct experimental outcomes which taken together in a complementary way form a complete quantum description.'¹ Of course, it is implicit that this principle is to be understood with reference to the breakdown of classical mechanics, which is characterised by the simultaneous applicability of the kinematical description and the dynamical description. Since in actual fact the experimental conditions required for a precise kinematical description are incompatible with those needed for a precise dynamical description, inferences of unlimited precision according to the classical rule (C. 2) are not possible. The limits of precision are given by Heisenberg's relations.

According to Schumacher, Bohr treats the quantum description as a *metalanguage* relative to the classical language, which describes the observational situation.² The appearance of statistical fluctuations reflects the impossibility of drawing precise inferences because of the incompatibility of the kinematical and dynamical descriptions needed for such inferences. The quantum metalanguage makes statements about the kinematical and dynamical descriptions, i.e. it involves rules of composition for kinematical and dynamical statements. These rules are, therefore, *necessarily statistical*. For Bohr, the 'quantum system' is an abstraction from the total ensemble of individual (classically describable) events in which it 'participates'. But this 'participation' cannot be analysed in terms of classical concepts—it is not the participation of a subsystem in an interaction. Rather, the relationship between the quantum system and the ensemble is like the relationship between a newspaper picture and the dots of which it is composed: the relationship is inherently statistical. In this sense, the quantum system belongs only to the ensemble.

3 THE HIDDEN VARIABLE APPROACH

Bohr's views form the basis of the Copenhagen interpretation of the quantum theory. The unsatisfactory feature of this interpretation is the

¹ D. L. Schumacher, 'Time and Physical Language'. Article in *The Nature of Time*, edited by T. Gold with the assistance of D. L. Schumacher (Ithaca, New York : Cornell University Press, 1967), p. 206.

² Ibid. p. 211.

'conservative thesis'; the assumption that unambiguously communicable information about what is observed is necessarily described in terms of the concepts of classical physics, or refinements of classical concepts. Here we need only be concerned with the question of whether this position is inevitable or not, without attempting to refute it on philosophical grounds. Evidently, the assumption of a necessarily classical language for the unambiguous description of the observational situation may in fact be unfounded. If so, it would be a very restrictive and stultifying assumption indeed, which could prevent the development of new kinds of unambiguous concepts that are potentially of experimental significance. If we reject this restriction, which is involved in Bohr's explanation of the distinction between quantum system and macroscopic measuring instrument, then the indeterminacy relations remain unexplained and the quantum measurement process is problematic. In order to formulate more precisely this measurement problem, as well as the hidden variable approach to its solution, we shall examine briefly the structure of the quantum theory.

The quantum theory may be characterized by the following features:

(Q. 1) A system is represented by a vector $|\psi\rangle$ in a Hilbert space. According to the Copenhagen interpretation, the significance of the vector $|\psi\rangle$ is that it may be used to generate sets of numbers (according to certain rules) which refer to the probabilities of possible measurement results.

(Q. 2) A physical quantity or 'observable' (R) is associated with a Hermitian, hypermaximal operator ($\mathbf{R}$) with a complete set of eigenvectors ($|R_1\rangle$, $|R_2\rangle$, ..., in the discrete case). The possible results of measurement of the observable R are restricted to the eigenvalues ($R_1, R_2, \dots$) of the associated operator $\mathbf{R}$.

(Q. 3) Two different rules define the change in $|\psi\rangle$ for any specified system and any specified initial $|\psi\rangle$:

(A) The rule which is applicable provided the system is not measured. One expression of this rule is Schrodinger's equation (a deterministic equation for the time evolution of $|\psi\rangle$).

(B) The rule which is applicable only when the system is measured. If the observable R is measured, then the probability of a transition $|\psi\rangle \rightarrow |R_i\rangle$ is $|\langle R_i|\psi\rangle|^2$. ($|\langle R_i|\psi\rangle|^2$ is a number equal to the magnitude of the square of the projection of the vector $|\psi\rangle$ onto the vector $|R_i\rangle$, the eigenvector of $\mathbf{R}$ corresponding to the eigenvalue R_i .) If $|\psi\rangle$ is projected or 'collapses' on to the vector $|R_i\rangle$ in a measurement of the observable R , then the result of the measurement is taken as R_i .

According to the Copenhagen interpretation, the 'measurement results' referred to in (Q. 1), (Q. 2), (Q. 3) are *classically describable*. (Q. 1) relates

the quantum description to the unambiguously definable *classical possibilities*. (Q. 2) states that these possibilities cannot all be realised simultaneously, i.e. possibilities associated with non-commuting operators are incompatible. The principle of complementarity relates the incompatibility of different kinds of definite results (associated with non-commuting operators in the algorithm) to the incompatibility of the experimental conditions needed to observe these results. Following Schumacher, the quantum postulates (Q. 1), (Q. 2), and (Q. 3) may be interpreted as metalinguistic statements. From this point of view the rule (Q. 3) which defines the change in $|\psi\rangle$ comprises a 'dynamical' rule (A) and a purely 'linguistic' rule (B). (A) refers to a relationship between a quantum metalinguistic concept and time. (B) refers to the relationship between the quantum metalanguage and the classical language, i.e. it asserts the essentially statistical relationship between the metalanguage of Hilbert space and the language of directly observable results.

The measurement problem is associated with the separate statistical rule (B) for measurement processes, which is a new feature relative to classical physical theories. The Copenhagen interpretation provides a consistent interpretation of the quantum theory, and thereby explains rule (B), but this explanation involves the assumption that the structure and function of the measuring instrument at the macro-level is restricted to what can be described in terms of the classical language. The hidden variable approach to the measurement problem, which avoids this assumption, leads to a new theory with new experimental consequences, in which the *order* of results in a sequence of measurements plays an important role in the theory. This is a non-classical concept, referring to unambiguously communicable observational information in Bohr's sense. (See section 5.)

The hidden variable approach provides a solution to the measurement problem in the following way. Firstly, it is proposed that a complete theory involves additional hidden variables $[\lambda]$, so that a system is represented by a set of variables $\{|\psi\rangle, [\lambda]\}$. (The notation $[\lambda]$ is used here to leave open for the moment the question of whether the additional variables $[\lambda]$ form a scalar, vector, etc.) These hidden variables determine the actual result in each case; averaging over the hidden variables yields the statistics. Secondly, it is proposed that there is a single rule for all changes in $\{|\psi\rangle, [\lambda]\}$, rather than two rules, one 'dynamic' and the other 'linguistic'. In addition, the hidden variable approach rejects the classical idea that what can be observed in a measurement process is in principle separable from the total set of experimental conditions, or of the environmental conditions in general, whether part of an experiment or not. Thus, it is not assumed that the environment relative to which a certain ('elementary') process occurs is

necessarily reducible to a mere complex of similar events, or that the environment influences the process merely through appropriate boundary conditions for the equations of motion. •

This means that the rule defining the change in $\{|\psi\rangle, [\lambda]\}$ is neither 'dynamical' nor 'linguistic'. It is a new kind of rule, with some features of both, but equivalent to neither. It is similar to a dynamical rule, in that it involves an equation of motion which determines the change in $\{|\psi\rangle, [\lambda]\}$ in each case, so that the quantum statistics is recovered by averaging over the hidden variables $[\lambda]$. It is 'trans-dynamical', in the sense that this equation of motion depends on the whole environment, including the measuring instrument as an aspect of this environment. The idea is in a way similar to Mach's principle, which relates the inertial properties of a micro-system (i.e. the differential equations of motion of a free particle) to the cosmological environment (the distribution of mass-energy in the universe). However, instead of involving only the cosmological environment, the hidden variable equation of motion depends on the whole environment on many levels, including the level of laboratory apparatus (i.e. the macro-environment). Since it is implicit that there is no fundamental small-scale level to which the large-scale levels are completely reducible, this hidden variable equation of motion does not in any sense refer to a 'rockbottom' level.

For example, consider a two-dimensional Hilbert space for simplicity. Measurement of the observable R with eigenvalues R_1, R_2 and eigenvectors $|R_1\rangle, |R_2\rangle$, means that either the transition $|\psi\rangle \rightarrow |R_1\rangle$ occurs or the transition $|\psi\rangle \rightarrow |R_2\rangle$ occurs (according to (B)). All that is required, in order to recover the quantum statistics on the average, is a suitable equation of motion involving $\{|\psi\rangle, [\lambda]\}$ in such a way that the transition $|\psi\rangle \rightarrow |R_1\rangle$ occurs for certain initial values of $[\lambda]$ and the transition $|\psi\rangle \rightarrow |R_2\rangle$ occurs for certain other values of $[\lambda]$. These values of $[\lambda]$ must be such that averaging over them will reproduce the statistics defined by $|\psi\rangle$ alone according to the algorithm of the quantum theory, i.e. $\text{prob}(R_1) = |\langle R_1|\psi\rangle|^2$, $\text{prob}(R_2) = |\langle R_2|\psi\rangle|^2$, $\text{Av}(R) = |\langle R_1|\psi\rangle|^2 R_1 + |\langle R_2|\psi\rangle|^2 R_2$.

It is important to realise here that the proposed rule of motion can only define the change in $\{|\psi\rangle, [\lambda]\}$ for *specific measurement processes* or, more generally, for *processes occurring relative to specified environments*. If a set of possible final vectors is defined (the set of eigenvectors of the observable measured), then the initial values of the variables $\{|\psi\rangle, [\lambda]\}$ determine a *particular member* of this set according to the associated rule. The *particular set* of possible final vectors is not defined by the initial values of the variables $\{|\psi\rangle, [\lambda]\}$, but by the *measuring instrument*. In other words, there is really a set of related rules, each individual rule defining the behaviour

of the system relative to a specific environment, which constitutes the measurement of a specific observable. In the hidden variable theory referred to¹ (which may be regarded as a particular model of the general hidden variable approach discussed here), it is assumed that the motion of an 'isolated' system is approximately described by Schrodinger's equation. But in any environment, including that required for a measurement, there are additional terms in the equation of motion. This modified Schrodinger equation defines a rule which relates a definite final state to each initial state in a way that depends on the macro-environment.

The essence of the Copenhagen interpretation is that the change in $|\psi\rangle$ during a measurement process cannot be understood as the result of a mechanical interaction between the measuring instrument and the system measured. Bohr argued that this fundamental indivisibility of the relationship between the measuring instrument and the system measured 'implies that a comprehensive account of experience in this field must be expressed by statistical laws'. Now, if hidden variables can be introduced in the manner outlined above, and if a viable theory can be developed with suitable equations of motion, then Bohr's argument for a necessarily statistical account of the quantum measurement process becomes questionable. For the introduction of additional hidden variables makes the change in $|\psi\rangle$ in principle predictable *for each specified measurement process*, i.e. the change in $|\psi\rangle$ is in principle predictable *even if the theory is not a mechanical theory designed to avoid the 'feature of wholeness'*.

The position adopted here is that the questionable part of Bohr's argument is the assumption that the structure and function of the measuring instrument at the macro-level is necessarily described in terms of classical concepts. It is the restriction of the macro-level to what can be described *classically* that is rejected in the hidden variable approach, and not the supposition of a fundamental distinction between the macroscopic measuring instruments and the lower level measured systems. With the hypothesis of hidden variables, this distinction is proposed as a new ontological thesis, i.e. a thesis about the nature of physical system and physical change.

It is proposed that the structure and function of any physical system and the structure and function of the environment of the system are internally related in a reciprocal way. In other words, what a system is structurally and how the system behaves is internally related to the (structural) constitution and (functional) behaviour of the environment of that system. No restrictive assumption is made about the language (classical or otherwise) which is used to describe the structure and function of the environment or the system.

¹ D. Bohm and J. Bub, op. cit.

On this view, a measurement process is merely a particular case of the universal *structural process*¹ (the mutual functioning of system and environment as internally related structures) that may be characterized as *reflective*,² in the sense that certain features of the behaviour of the environment and certain features of the behaviour of the system 'mirror' each other. Thus, a measuring instrument is not conceived as a system which interacts mechanically with another system, revealing by a certain manifest change a particular property or attribute which can then be predicated of the system measured (as in classical mechanics). What there is, is a total hierarchical (i.e. multi-level) structural process of a certain kind. A measuring instrument may be abstracted from this structural process at a certain level in the sense that a set of possible reflective functions may be specified in relation to certain reflected functions of a 'measured' structure. For example, if a set of eigenvectors to which a transition is possible is specified, this amounts to abstracting a measuring instrument and measured system from a structural process, which is thereby interpreted as the 'measurement of an observable'. But there is no basic set of variables which can be associated with the abstracted systems in such a way that the observables are simply functions of these variables, i.e. variable kinematical or dynamical quantities referring to the separate systems (as in classical mechanics). Moreover, the fundamental distinction between the measuring instrument and the system measured does not lie in the difference between classical statements about the macro-level and quantum metalinguistic statements, from which the micro-level is abstracted (as in the Copenhagen interpretation of the quantum theory).

In the hidden variable approach, both macro-level and micro-level are regarded as *aspects* of the whole structural process, i.e. structural-functional abstractions from the whole. The equation of motion referred to relates the functioning of the whole structural process at the micro-level to certain features of this structural process at the macro-level, i.e. to the environment at the level of the laboratory. So, in a certain way, this equation incorporates a structural relationship similar to that of the relationship between the quantum metalanguage and the classical language, in Bohr's interpretation of the rule (B). However, we are now concerned with an ontological structure, not a linguistic structure. Because the laws of the micro-level ultimately depend on *all the levels*, these laws transcend the concept of dynamics, as it is usually understood. They involve a conception of wholeness,

¹ The concept of a 'structural process' is developed by D. Bohm in 'Space, Time and the Quantum Theory—Understood in terms of Discrete Structural Process', *Proceedings of the International Conference on Elementary Particles* (Kyoto: 1965).

² Ibid. The concept of a 'reflective function' is developed by Bohm in this article.

but this conception has nothing to do with the notion that only classical concepts can capture the unambiguously communicable features of the observational situation. Rather, the theory makes use of new concepts, such as $|\psi\rangle, [\lambda]$, the state of the macro-environment, etc., and as shown in section 5 the possibility of new experimental results is implicit in these concepts. They may be regarded as a first step towards the development of an ordered hierarchy of inter-related concepts referring to all levels of the total structural process (e.g. the concepts of relation, order, structure, function, reflective function, process, etc.).¹

4 STANDARD ARGUMENTS AGAINST THE HIDDEN VARIABLE HYPOTHESIS

A number of objections have been raised against the hidden variable strategy in general. The 'common sense' or traditional argument is a naive objection based on a misunderstanding of the hidden variable programme, which is seen as an attempt to establish quantum theory as the statistical mechanics of an essentially *classical* subquantum level. Thus, it is assumed that the observables R, S , etc., are treated as functions of the 'state variables' $\{|\psi\rangle, [\lambda]\}$ in the theory. In this case, if $R = R(|\psi\rangle, [\lambda])$ and $S = S(|\psi\rangle, [\lambda])$, $R' + S' = (R + S)'$ for any two observables R, S . (The " ' " denotes a particular value of the observable for particular values of $|\psi\rangle$ and $[\lambda]$; R, S , $(R + S)$ are defined as the observables associated with the Hermitian operators $\mathbf{R}, \mathbf{S}$, and $(\mathbf{R} + \mathbf{S})$ respectively.) It is pointed out that such a hidden variable theory can be ruled out immediately, since the eigenvalues of non-commuting operators are in general not additive.

This argument, in one form or another, has been made over and over again in the historical development of the quantum theory, so that there is still a very widespread belief that it remains the chief obstacle to any hidden variable theory, if not an insurmountable obstacle. A clear and explicit formulation is contained in the classic *La Théorie de l'Observation en Mécanique Quantique*, by F. London and E. Bauer.² In section 8 of their book, London and Bauer demonstrate that a 'pure' ensemble (represented in the quantum theory by a projection operator) is homogeneous in the sense that it does not contain sub-ensembles with different statistical properties. Of course, there is still the question of whether the pure ensembles could be decomposed into sub-ensembles in some other way, i.e. whether they could be represented as mixed ensembles of systems 'bien

¹ Ibid. The process metaphysics involved here is close to the ideas of A. N. Whitehead in *Process and Reality* (New York: Macmillan, 1929).

² F. London and E. Bauer, *La Théorie de l'Observation en Mécanique Quantique* (Paris: Hermann et Cie, 1939).

définis au sens ordinaire de la physique classique'.¹ London and Bauer consider the following example² to show that this is in fact impossible. For a particle of spin $\frac{1}{2}$, there are only two possibilities for the value of the spin in any given direction: $+\frac{1}{2}$ or $-\frac{1}{2}$. Let $|S_+\rangle$ and $|S_-\rangle$ be the eigenvectors associated with the two possible values of the component of the spin (S) along the polar axis of a spherical co-ordinate system. Suppose the system is represented initially by the Hilbert space vector $|\psi\rangle = |S_+\rangle$. Then the probabilities for the two possible spin values are: $P_+ = 1$, $P_- = 0$. The only possible values for the component of the spin in some other direction (S'), defined by the angles θ , ϕ , are again $+\frac{1}{2}$ and $-\frac{1}{2}$. $|\psi\rangle$ also defines the probabilities for the two possible spin values in this direction. If the corresponding eigenvectors are $|S'_+\rangle$ and $|S'_-\rangle$, then:

$$|\psi\rangle = |S_+\rangle = c_+|S'_+\rangle + c_-|S'_-\rangle \quad (4.1)$$

with

$$c_+ = e^{i\phi/2} \cos \theta/2, \quad c_- = e^{-i\phi/2} \sin \theta/2 \quad (4.2)$$

and

$$P'_+ = |c_+|^2 = \cos^2 \theta/2, \quad P'_- = |c_-|^2 = \sin^2 \theta/2 \quad (4.3)$$

London and Bauer argue:

Il est évident qu'il n'est pas possible de décomposer cette statistique en un mélange de spins bien orientés: il faudrait dans ce mélange qu'une fraction $\cos^2 \theta/2$ des atomes soit en direction θ , ϕ avec une composante $1/2 \hbar/2\pi$, une fraction $\sin^2 \theta/2$ en direction opposée, ce qui serait encore possible en soi. Mais on devrait avoir en outre 100% des atomes avec la même composante $1/2 \hbar/2\pi$ dans la direction de l'axe polaire et enfin une fraction $\cos^2 \theta'/2$ suivant toute autre direction θ' . Ce serait là un tour de passe-passe bien difficile à réaliser! Il est évidemment impossible de former un ensemble virtuel d'atomes orientés qui comporte en même temps toutes ces statistiques. Déjà leur forme mathématique exclut cette possibilité.³

This whole analysis is beautifully simple, perfectly correct, but quite irrelevant to the hidden variable approach. Given the initial values of the Hilbert space vector and the hidden variables, then within the framework of a hidden variable theory it is possible to predict whether the final value of the Hilbert space vector will be $|S_+\rangle$ or $|S_-\rangle$ if the reflective structural process from which the system is abstracted is such as to represent a measurement of the observable S , and it is equally possible to predict whether the final value of the Hilbert space vector will be $|S'_+\rangle$ or $|S'_-\rangle$ if the observable S' is measured. The hidden variables do not decompose the quantum ensemble into sub-ensembles which consist of systems with definite spins in the classical sense, i.e. in which each system is associated with a unique spin vector *defined simply as a function of the hidden variables*.

¹ F. London and E. Bauer, *La Théorie de l'Observation en Mécanique Quantique*, p. 32.

² Ibid. pp. 32, 33.

³ Ibid. pp. 33, 34.

The aim of a hidden variable theory is rather to re-interpret the notion 'measurement of an observable' in a non-classical, non-Copenhagen sense.

The argument of London and Bauer is essentially the same as the argument against the hidden variable hypothesis based on the two slit thought experiment. An ensemble of electrons passing through a two slit system produces interference fringes on a photographic plate. It is often pointed out that the phenomenon cannot be understood as the motion of a cloud of classical-like particles, each of which actually passes through a definite slit, for then the intensity at any point of the plate (i.e. the number of impinging particles) would be merely the sum of the intensities from both slits. This is incompatible with the appearance on the plate of dark areas which would be light with one slit closed. To maintain such a classical theory, it would be necessary to assume that the mere existence of an open slit, through which the electron particle did not pass, could inhibit the passage of the particle to an area on the plate where it would have arrived had the slit been closed. Since the same interference pattern is observed when the electrons proceed through the slit system one at a time, it is not possible to explain this result by some sort of interaction between the particles which pass through the different slits.

The two-slit experiment is used as an argument against hidden variables in the following way. With only slit *A* open, it should be possible within the framework of a hidden variable theory to predict whether or not the particle will pass through the slit, given the values of the relevant hidden variables. Similarly, with only slit *B* open, it should be possible to predict whether or not the particle will pass through slit *B*. Therefore, with both slits open, it should be possible to predict whether an individual electron will pass through slit *A* or slit *B*. But then the different values of the hidden variables would effect a decomposition of the ensemble into particles which will pass through slit *A* and particles which will pass through slit *B*, and such a decomposition is incompatible with the appearance of interference fringes.

The same limited conception of the hidden variable hypothesis is involved here as in the London and Bauer argument. Even if the values of the hidden variables and the Hilbert space vector of the system are known, nothing can be predicted without reference to a specific environment, without which there is no reflective structural process. Given the values of the hidden variables and a one slit system, it should indeed be possible to predict whether or not a particular electron will pass through the slit, but only in the sense that its presence will or will not be registered at any chosen point on the photographic plate. Given the values of the hidden variables and a two slit system, it should again be possible to predict whether or not

the presence of a particular electron will be registered at any chosen point on the photographic plate, but the possibility of this prediction is evidently neutral with respect to the question of whether or not the electron 'passed through' a particular slit. This question refers to a different environment with its own associated possibilities.¹

A more subtle objection to a hidden variable theory of the kind envisaged here can be formulated as a corollary to von Neumann's hidden variable proof.² Within a certain axiomatic framework which essentially defines the statistical relations of the quantum theory, von Neumann was able to represent every ensemble by a statistical operator. Since within this framework there exist no physically meaningful dispersion-free ensembles, while there do exist pure ensembles which are homogeneous in the sense that they contain no sub-ensembles with different statistical properties, von Neumann concluded that the statistics of the pure ensemble could not result from averaging over hidden variables.

Now, one of the axioms necessary to the conclusion that all allowable ensembles can be represented by a statistical operator is the 'additivity postulate'.³ Von Neumann suggested that any suitable theory would at least have to require that $Av(R) + Av(S) = Av(R+S)$ for any observables R, S in order to reproduce the quantum statistics, since this relation holds for quantum ensembles. To understand the significance of this requirement for a hidden variable theory constructed along the lines suggested, it is necessary to be a little more specific. Let $[\lambda] = \lambda$ be a scalar which can vary between 0 and 1. The quantum statistics can be reproduced, for example, by an equation of motion for which $|\psi\rangle \rightarrow |R_1\rangle$ if $0 \leq \lambda \leq |\langle R_1 | \psi \rangle|^2$ and $|\psi\rangle \rightarrow |R_2\rangle$ if $|\langle R_1 | \psi \rangle|^2 < \lambda \leq 1$, since then:

$$Av(R) = \int_0^{|\langle R_1 | \psi \rangle|^2} R_1 d\lambda + \int_{|\langle R_1 | \psi \rangle|^2}^1 R_2 d\lambda = |\langle R_1 | \psi \rangle|^2 R_1 + |\langle R_2 | \psi \rangle|^2 R_2 \quad (4.4)$$

¹ This experiment is analysed somewhat differently in the earlier hidden variable theories which developed from de Broglie's original idea of a 'pilot wave'. See, for example, D. Bohm *Causality and Chance in Modern Physics* (New York: Van Nostrand Company, Inc., 1957), pp. 111-16. In these theories, the electron is regarded as a combination of particle and ψ field which undergoes suitable random fluctuations in its motion. The form of the ψ field (and hence the hidden variables which describe the fluctuations) depends on the macro-environment. But since the particle is assumed to have a definite position correlated to the values of these variables, it is evidently possible in principle to describe the entire trajectory of the particle as it passes through the slits. The type of theory discussed above involves a much more radical departure from the classical wave-particle dichotomy, although even the early hidden variable theories are non-classical as far as the measurement process is concerned. However, the 'feature of wholeness' is rather obscure in these theories.

² J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton: Princeton University Press, 1955). Translated from the German by R. T. Beyer. Chapter 4, Parts 1 and 2, pp. 295-328.

³ The significance of this assumption was first pointed out by J. S. Bell, *Rev. Mod. Phys.* 38 (1966), 447.

if λ is assumed to be completely random. Clearly, in this case, $Av(R) + Av(S) = Av(R+S)$. But the measurement of R partially determines λ , which can be restricted either to the interval $0 \leq \lambda \leq |\langle R_1 | \psi \rangle|^2$ or to the interval $|\langle R_1 | \psi \rangle|^2 < \lambda \leq 1$ immediately after the measurement. Hence the quantum statistics will not be reproduced for the results of measurements of an observable S (where S does not commute with R) on a representative ensemble of *such* systems—i.e. systems for which the value of λ has been partially determined by a measurement of the observable R . Thus, in general, there could exist ensembles violating the quantum statistics, and it can easily be shown that for such ensembles $Av(R) + Av(S) \neq Av(R+S)$ (where $R, S, (R+S)$ are now assumed to be measured on the systems of similar ensembles, prepared in each case by the measurement of some other observable T).

It follows, then, as a corollary to von Neumann's proof, that any hidden variable theory of this type will involve the existence of statistical ensembles which will not in general obey the additivity postulate. While London and Bauer consider only hidden variable theories which involve the relation $R' + S' = (R+S)'$ between observables and point out that this relation is in general false for quantum observables, the von Neumann proof can be interpreted as the statement that some statistical ensembles of a hidden variable theory will violate the relationship $Av(R) + Av(S) = Av(R+S)$, which is always true for quantum ensembles. (See Appendix II.) Both arguments impose restrictions on viable hidden variable theories. The London and Bauer argument immediately rules out theories based on the classical statistical mechanical model in an obvious or naive way (which have probably never been seriously considered anyway). The dilemma posed by the von Neumann proof is less trivial, but its solution clarifies an important aspect of the hidden variable hypothesis. One way out is to postulate that the hidden variables λ are effectively randomised in a long enough interval of time τ after the measurement of an observable. In this case, the quantum statistics becomes an approximation valid only under the condition that the time interval between successive measurements is greater than τ . (See section 5.)

Of course, von Neumann was not concerned with this possibility because he did not approach the problem in this way. His proof only attempts to demonstrate that within a certain axiomatic framework broad enough to include the quantum statistics there are no dispersion-free ensembles. In other words, the hidden variable hypothesis is incompatible with the set of axioms. The above argument shows how a particular type of hidden variable theory can conflict with a particular axiom in this set and yet incorporate the quantum statistics as a special case. Whether von Neumann would

have thought that his proof automatically implied a refutation of such a theory is a debatable point. It is clear, however, that many people have interpreted the proof in this way, probably as the result of von Neumann's concluding remarks that: '... we need not go any further into the mechanism of the 'hidden parameters' since we now know that the established results of quantum mechanics can never be re-derived with their help'¹ and that: '... it is therefore not, as is often assumed, a question of a re-interpretation of quantum mechanics,—the present system of quantum mechanics would have to be objectively false, in order that another description of the elementary processes than the statistical one be possible.'²

These statements are ambiguous. They can be interpreted to mean only that the quantum theory cannot be reformulated in terms of hidden variables without involving the possibility of more general statistical relations, i.e. that the theory is in principle *falsifiable*. But this is a trivial and almost obvious fact. Perhaps this is all that von Neumann wished to establish, but by a mathematically rigorous argument. On the other hand, von Neumann's remarks can be interpreted as the assertion that the quantum theory would already have been *falsified* by the current experimental results, if a hidden variable theory were possible. The fact that the quantum theory has not been falsified would then mean that any conceivable hidden variable theory would be falsified by the current experimental results. Evidently there could never be any satisfactory proof of such a position, and it is probably unfair to attribute this interpretation to von Neumann. However, the psychological effect of the proof on subsequent research can definitely be associated with this interpretation.

There are further objections to the rather oversimplified hidden variable scheme introduced above, but these can all be overcome by choosing $[\lambda] = \langle \lambda |$, a vector in the dual Hilbert space of $|\psi\rangle$, formulating explicit equations of motion, etc.³ The time evolution of $|\psi\rangle$ is described by a modified Schrodinger equation, in which an additional term (regarded as negligible for an 'isolated' system) explicitly effects the projection of $|\psi\rangle$ on to an eigenvector during a measurement process. The eigenvector is determined in each individual case by the relationship between $|\psi\rangle$ and $\langle \lambda |$. It becomes clear in such an analysis that, implicitly or explicitly, the objections to hidden variables can be traced to the unjustified assumption that a hidden variable theory is an attempt to reconstruct the quantum statistics in order to avoid the essential 'feature of wholeness' of quantum processes. This assumption is supplemented by the false corollary that 'wholeness' is associated with 'rockbottom indeterminism', or the

¹ J. von Neumann, op. cit. p. 324.

² Ibid. p. 325.

³ D. Bohm and J. Bub, op. cit.

'irreducibility' of the quantum statistics. Hence the hidden variable approach is regarded as a reactionary and fruitless attempt to return to a Newtonian-type determinism. •

5 EXPERIMENTAL CONSEQUENCES OF THE HIDDEN VARIABLE HYPOTHESIS

The practical difference between the Copenhagen interpretation and the hidden variable approach suggested here can be made clear by a 2-dimensional example: the theory of the spin- $\frac{1}{2}$ particle.

With respect to the representation defined by the spin observable S , the Hilbert space vector of the system is $|\psi\rangle = \psi_1|S_1\rangle + \psi_2|S_2\rangle$. The hidden variables are represented by a vector in the dual Hilbert space: $\langle\lambda| = \lambda_1\langle S_1| + \lambda_2\langle S_2|$. The equations of motion for the system in an 'S-environment' are postulated as:

$$\begin{aligned}\frac{d\psi_1}{dt} &= \Gamma_1(\psi_1, \psi_2) + \Delta_1(\psi_1, \psi_2, \lambda_1, \lambda_2) \\ \frac{d\psi_2}{dt} &= \Gamma_2(\psi_1, \psi_2) + \Delta_2(\psi_1, \psi_2, \lambda_1, \lambda_2)\end{aligned}\quad (5.1)$$

Γ_1 and Γ_2 represent the Schrodinger change. In matrix form in this representation, the Schrodinger equation is:

$$i\hbar \frac{d}{dt} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \quad (5.2)$$

so that $\Gamma_1 = -i/\hbar(H_{11}\psi_1 + H_{12}\psi_2)$, and $\Gamma_2 = -i/\hbar(H_{21}\psi_1 + H_{22}\psi_2)$. The Δ terms are non-linear functions of $\psi_1, \psi_2, \lambda_1, \lambda_2$.¹ These terms may be regarded as negligible, except during a measurement process, when their effect is much greater than that of the Schrodinger terms.

The equations:

$$\begin{aligned}\frac{d\psi_1}{dt} &= \Delta_1(\psi_1, \psi_2, \lambda_1, \lambda_2) \\ \frac{d\psi_2}{dt} &= \Delta_2(\psi_1, \psi_2, \lambda_1, \lambda_2)\end{aligned}\quad (5.3)$$

define a transition from $|\psi\rangle$ to the eigenvector $|S_1\rangle$ or to the eigenvector $|S_2\rangle$, depending on the initial values of the variables $\psi_1, \psi_2, \lambda_1, \lambda_2$. It can be shown² that the quantum statistics can be recovered from a random distribution of these variables over the hypersphere of unit radius in the dual Hilbert space, defined by $|\lambda_1|^2 + |\lambda_2|^2 = 1$. However, the effect of a specific measurement process is to restrict the values of the hidden variables

¹ See *ibid.* p. 463 for an explicit form of these functions.

² *Ibid.* pp. 463, 464.

to a certain domain of the unit hypersphere (compatible either with the transition $|\psi\rangle \rightarrow |S_1\rangle$ or with the transition $|\psi\rangle \rightarrow |S_2\rangle$). Therefore, the quantum statistics will not be reproduced in general for a subsequent measurement of the spin component in some other direction, unless the λ_1, λ_2 variables undergo a quasi-ergodic movement (negligible during the measurement process), leading to an eventual random distribution of these variables over the unit hypersphere. In a long enough time τ after a measurement the λ_1, λ_2 variables may be completely randomised, so that the theory would imply the usual quantum statistics for the results of a subsequent measurement.

This means that an ensemble of *ordered sequences of measurement processes* (involving observables with non-commuting operators) can in principle yield a distribution of results which contradicts the quantum statistics, if the measurement processes follow each other sufficiently rapidly.¹ In this way the hidden variables become in principle measurable *on the basis of information derived from the order of results in a suitable sequence of measurements*. According to the Copenhagen interpretation, the order of results in a sequence of measurements has no theoretical significance. The probabilities (and *only* the probabilities) of possible measurement results at any time are specified by the Hilbert space vector representing the system at that time—the previous history of the system is irrelevant.

The relationship of the hidden variables to the order of events at the macro-level points to the real significance of the hidden variable hypothesis. It is a formal device for introducing the idea that the macro-order is not necessarily randomly related to the order of functioning at the micro-level, without thereby implying that the macro-level is reducible to the micro-level in a mechanical sense. The hidden variable theory associates a set of potentialities with each macroscopic experimental arrangement, symbolized by the set of eigenvectors of the observable measured by the apparatus. The variables $\psi_1, \psi_2, \lambda_1, \lambda_2$, which reflect the micro-order, define the propensity² to realise any given potentiality in an actual experiment. If it is assumed that the hidden variables λ_1, λ_2 are always random, then the theory is equivalent to the orthodox quantum theory, and the propensities are restricted to the quantum probabilities. Thus, in the orthodox quantum theory, the micro-order as represented by the variables ψ_1, ψ_2 is randomly related to the order of events at the macro-level, in the sense that the

¹ An experimental comparison between the statistical predictions of the quantum theory and this hidden variable theory has been carried out recently. On the basis of a photon polarisation experiment, it has been shown that if the hidden variable hypothesis is true, then $\tau < 2.4 \times 10^{-14}$. See Costas Papaliolios, *Phys. Rev. Letters*, **18** (1967), 622.

² The term 'propensity' in the context of the quantum theory is due to K. R. Popper. See 'Quantum Mechanics without "The Observer"', article in *Quantum Theory and Reality*, ed. Mario Bunge (New York: Springer-Verlag, 1967).

relationship of ψ_1, ψ_2 to the potentialities defined by a macroscopic experimental arrangement is inherently statistical. In the hidden variable theory, the λ_1, λ_2 variables may have any distribution over the surface of the unit hypersphere, from a delta function distribution to a random (uniform) distribution. This possibility means that the propensity to realise a given potentiality depends on the total set of experimental conditions, including previous experimental results, in such a way that the propensity can have any value between 0 and 1 (including 0 and 1) in a situation in which the relationship between the quantum state ψ_1, ψ_2 and the outcome of the measurement is inherently statistical.

The very possibility of such a theory refutes Bohr's thesis that unambiguously communicable information about what we observe can only be captured by classical concepts, or some refinement of classical concepts. The concept of the order of results in a sequence of measurement processes ('unambiguously communicable information') belongs neither to the kinematical concepts nor to the dynamical concepts of the classical description. Here we see the crucial importance of considering alternative axiomatic and conceptual structures *even before the old theories have been refuted*. For these new conceptual structures suggest new experiments, i.e. new ways to falsify current theories, that are not likely to be considered otherwise. As Popper has repeatedly emphasised, and more recently Feyerabend,¹ the empirical content of our knowledge is always improved by the existence of a counter-theory. The counter-theory shows what it would *mean* to falsify the current theory. Either the current theory is falsified by experiment, or it is not. In both cases we learn something new about the domain of validity of the current theory. As Feyerabend aptly puts it: 'Both the relevance and the refuting character of many decisive facts can be established only with the help of other theories that, although factually adequate, are not in agreement with the view to be tested. This being the case, the production of genuine refutations may have to be preceded by the invention and articulation of alternatives to that view'.² The failure to consider alternatives leads to dogmatism, because the current theory always seems inevitable in the absence of an awareness of the alternative possibilities.

6 CONCLUSION

The central thesis of this article is that a hidden variable theory of quantum phenomena is not necessarily incompatible with Bohr's 'feature of

¹ P. K. Feyerabend, 'Problems of Empiricism'; article in *Beyond the Edge of Certainty: Essays in Contemporary Science and Philosophy*, ed. R. G. Colodny (Englewood Cliffs, New Jersey: Prentice-Hall, 1965).

² Ibid. p. 176.

wholeness'. A hidden variable theory can be developed in which the quantum theory is explicitly not regarded as the statistical mechanics of an essentially classical subquantum level. The theory involves testable consequences which contradict orthodox quantum theory. In addition, the theory introduces new concepts which seem interesting and worth further study. The basic idea is that of a multi-level theory, each level being an aspect of (or abstraction from) a universal structural process. Each aspect of the structural process partially reflects every other aspect. Observation or measurement is associated with this reflective process. This implies that what is observed cannot be regarded as 'existing out there' apart from the process in which it can be observed.

From this point of view, the Copenhagen interpretation of the quantum theory is characterised by two theses: a 'revolutionary thesis' and a 'conservative thesis'. The revolutionary thesis, that the macroscopic measuring instruments and microscopic measured systems are on different levels of functioning which are not related mechanically, is developed further by the hidden variable approach. The conservative thesis, that unambiguous information about what is observed is necessarily described in terms of the concepts of classical physics, is rejected. The hypothesis of hidden variables is a formal device for introducing the idea that the relationship between different levels of functioning is potentially 'creative'.¹ The order of events at the macro-level is not completely reducible to the micro-order in a mechanical sense, but the micro-order is related to the propensity to realise a given potentiality at the macro-level. This propensity is not completely defined by the relationship between two classically describable experimental arrangements (the first arrangement preparing the initial quantum state and the second arrangement performing the measurement), but depends on a much wider class of environmental conditions, in such a way that the order of previous experimental outcomes can influence the propensity. The relationship between different levels of functioning of the structural process is neither mechanically reducible to the deterministic functioning of a 'rockbottom level', nor inherently statistical. It is 'creative' in the sense that, under suitable circumstances, a 'random' ordering of events may become 'less random' and, in the limiting case, 'deterministic'. In other words, new orders may be originated and old orders annihilated.

Thus, the deep intention behind the hidden variable approach is not in conflict with the essence of Bohr's interpretation of the quantum theory. The hidden variable hypothesis may be regarded as a stage in the transition

¹ For a discussion of this notion of 'creativity', see D. Bohm, 'Space, Time and the Quantum—Understood in terms of Discrete Structural Process', *Proceedings of the International Conference on Elementary Particles* (Kyoto, 1965).

from the Copenhagen interpretation to a structural process physics, which will involve Bohr's idea of 'wholeness' in a new and essential way.

APPENDIX I

One expression of the fundamental dynamical rule is given by Hamilton's equations of motion:

$$\begin{aligned}\frac{dq_i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial q_i}\end{aligned}\tag{AI.1}$$

$H(q_i, p_i)$ is the Hamiltonian function characteristic of the specific system. Suppose that two systems $S_1: \{q_1, p_1\}$ and $S_2: \{q_2, p_2\}$ interact for a short time and then move independently. If the interaction Hamiltonian during this time is, for example, $H_I = \alpha(q_1 p_2 + q_2 p_1)$, then the solutions to Hamilton's equations of motion are:

$$\begin{aligned}q_1 &= Ae^{\alpha t} + Be^{-\alpha t} & q_2 &= Ae^{\alpha t} - Be^{-\alpha t} \\ p_1 &= Ce^{\alpha t} + De^{-\alpha t} & p_2 &= -Ce^{\alpha t} + De^{-\alpha t}\end{aligned}\tag{AI.2}$$

where A, B, C, D are constants. (α is a constant which determines the strength of the interaction. It is assumed that α is large enough so that during the time of the interaction H_1 and H_2 —the Hamiltonians of the independent systems S_1 and S_2 —can be neglected in comparison with H_I .) Clearly, two observations of q_1 (or p_1) before and after the interaction are sufficient to determine the constants A and B (or C and D), and hence the values of q_2 (or p_2) before or after the interaction (or indeed at any time) can be deduced from these observations.

APPENDIX II

The difference between von Neumann's proof and the London and Bauer type of argument against hidden variable theories can be made clear by the analogy with statistical mechanics.

A classical dynamical variable is defined as a function of the coordinates and momenta (hidden variables) which determine the state of the system:

$$R = R(q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n)\tag{AII.1}$$

which is abbreviated below symbolically to $R(q, p)$. If only certain thermodynamic (i.e. large-scale) variables are measured, then the expectation value of R for the system is computed as the average over a suitably defined ensemble of systems:

$$\langle R \rangle = \int R(q, p) \rho(T, q, p) dq dp\tag{AII.2}$$

where T represents the thermodynamic variables and ρ is a probability density. In the equilibrium case, the statistical properties of the system are derived from the 'normal' Boltzmann distribution function

$$\rho_N = C \exp(-H(q,p)/kT),$$

where C and k are constants and T is the temperature. In general we may consider any distribution function $\rho(T,q,p)$, in particular the extreme case of a delta function, representing a dispersion-free ensemble in which the values of q,p are precisely defined and determine a value for every 'observable' $R(q,p)$ without dispersion.

Similarly, in a hidden variable theory of the quantum statistics, there should be 'non-normal' distributions of these variables possible, as well as a 'normal' distribution which would reproduce the quantum statistics. The extreme case of such a 'non-normal' distribution would be the dispersion-free ensemble, in which the values of all hidden variables are precisely defined and hence determine a particular value for every observable. Von Neumann demonstrated that the existence of such non-normal ensembles, with statistical properties different from those of the quantum theory (which is certainly implicit in any hidden variable theory), is impossible.

The relation between the average or expectation value of R , the matrix R_{mn} of the operator $\mathbf{R}$, and the statistical matrix U_{mn} of the statistical operator $\mathbf{U}$, is:

$$\langle R \rangle = \sum_{mn} U_{nm}(|\psi\rangle) R_{mn} = \text{Trace}(\mathbf{U}\mathbf{R}) \quad (\text{AII.3})$$

This is a *linear relation* between the expectation value of R and the matrix elements R_{mn} , i.e. *von Neumann's axioms involve the assumption that observable averages are linear functions of the operator matrix elements R_{mn} .*

London and Bauer, on the other hand, characterise hidden variable theories with the more restrictive requirement that *particular values of the observables are linear functions of the operator matrix elements R_{mn} :*

$$R = \sum_{mn} W_{nm}(|\psi\rangle, [\lambda]) R_{mn} \quad (\text{AII.4})$$

(This is evidently contradicted by the fact that the eigenvalues of observables with non-commuting operators are non-additive.) If this relation holds (i.e. if each system can be characterised by a $\mathbf{W}$ operator in this way), then:

$$\begin{aligned} \langle R \rangle &= \sum_{mn} \int W_{nm}(|\psi\rangle, [\lambda]) R_{mn} \rho([\lambda]) d[\lambda] \\ &= \sum_{mn} U_{nm} R_{mn} = \text{Trace}(\mathbf{U}\mathbf{R}) \end{aligned} \quad (\text{AII.5})$$

where:

$$U_{nm} = \int W_{nm}(|\psi\rangle, [\lambda]) \rho([\lambda]) d[\lambda] \quad (\text{AII.6})$$

Also:

$$\begin{aligned} \langle R+S \rangle &= \sum_{nm} \int W_{nm}(|\psi\rangle, [\lambda]) (R+S)_{mn} \rho([\lambda]) d[\lambda] \\ &= \sum_{nm} \int W_{nm} (R_{mn} + S_{mn}) \rho([\lambda]) d[\lambda] \\ &= \sum_{nm} \int W_{nm} R_{mn} \rho([\lambda]) d[\lambda] + \sum_{nm} \int W_{nm} S_{mn} \rho([\lambda]) d[\lambda] \\ &= \langle R \rangle + \langle S \rangle \end{aligned} \quad (\text{AII.7})$$

So the weaker von Neumann assumption follows from the London and Bauer assumption.

In a hidden variable theory, the value of R is determined by some *non-linear function* of $|\psi\rangle$, the hidden variables, and the matrix R_{mn} :

$$R = F(|\psi\rangle, [\lambda], R_{mn}) \quad (\text{AII.8})$$

In an ensemble:

$$\langle R \rangle = \int F(|\psi\rangle, [\lambda], R_{mn}) \rho([\lambda]) d[\lambda] \quad (\text{AII.9})$$

(i.e. observable averages are related in a non-linear way to the matrix elements R_{mn}), and the quantum statistics is obtained from a 'normal' ensemble specified by some $\rho_N([\lambda])$. Thus, the statistical matrix is no longer characteristic of all conceivable ensembles, and von Neumann's proof does not apply to this theory.

University of Minnesota

The Case Against Teleological Reductionism

by LARRY WRIGHT

I INTRODUCTION

Following the vitalism/mechanism controversy earlier this century, there seemed to prevail among empiricist philosophers an almost paranoid suspicion of teleological concepts. Anything that could cause so deep a confusion and so sharp a division among intelligent men must, it seemed, be intrinsically obfuscatory. Hence, wherever it appeared, the smoke of teleological terminology implied the fire of sloppy thinking.

Perhaps partly as an apology for this over-reaction, several empiricists in the past two or three decades have made attempts to show that the use of teleological concepts need not—at least not always and necessarily—indicate a logical or argumentative derangement. Although these attempts differ significantly in detail, they generally employ a reductionist strategy: they show that the teleological description of a phenomenon (or class of phenomena) is equivalent to some other description which employs only palpably respectable terms. Accordingly, the teleological description must have the same respectable logical status as the non-teleological one.

Implicitly, but often quite explicitly too, these arguments offer the equivalence of a respectable non-teleological expression as a necessary condition for the legitimacy of a teleological expression. This seemingly straightforward criterion plays a remarkably subtle and complex role in the major reductionist analyses of teleological concepts. In fact, the criterion itself can hardly be understood without examining its use in those analyses.

In this paper I critically examine the three classic reductionist analyses of teleology in recent philosophy. This examination leads to a detailed elaboration of the reductionist criterion, and eventually to a criticism of it as inadequate to the task of distinguishing legitimate teleological expressions from illegitimate ones.

2 LOGICAL AND METHODOLOGICAL PRELIMINARIES

The position that every legitimate teleological description is, or must be, equivalent to some non-teleological description of the same phenomenon

Received 18 April 1968

has been expressed in several different ways. The teleological account has been said to be 'translatable into', 'reducible to', or merely to be 'saying the same thing as' the non-teleological account. More elaborately, the point has been made by saying that the two accounts have the same 'cognitive content' and differ only in 'emphasis'. The equivalence between the two kinds of accounts being asserted by these expressions is a logical equivalence of some sort. Expressions like 'same cognitive content', 'translatable', 'saying the same thing' all concern the sense of the descriptions. Accordingly, the teleological description is to have the same sense as the non-teleological one. The two are logically equivalent descriptions.

Accordingly, the position I am considering in this paper holds not merely that whenever it is appropriate to describe a phenomenon teleologically it will *as a matter of empirical fact* always be the case that the same phenomenon can be described non-teleologically. The position holds that a legitimate teleological description *logically entails* a non-teleological one.

Moreover, if we are to maintain that the teleological account of a phenomenon is in any important sense 'reducible to', 'parasitic on', or 'dispensable in favour of', the non-teleological account, the equivalence of the two *must* involve their sense, their content. The position that teleological descriptions involve nothing new epistemologically requires that the two accounts be logically equivalent.

It is important to emphasize this because the argument of this paper cannot even begin until there is a distinction made between descriptions which are equivalent in this tight, logical sense and descriptions which might be corresponding or correlative in one way or another, but which, if equivalent at all, are so in a weaker sense. The phenomena which are usually the object of teleological description—behaviour patterns among living things, etc.—can all be described in terms of positions, colours, shapes and other purely physico-geometrical characteristics. A phenomenon as elaborate as the mating of Siamese fighting fish can be described as a *series* of events, without in any other way relating the individual events to one another, and without hinting that there might be other objects of similar appearance and behaviour. I might say: 'First this happened, then this, and then this'. And the demonstrative pronoun could, in each case, be replaced by a statement in terms of blobs of colour, more or less solid objects, motions of these objects, etc., containing no mention of mating, fish or even life. But the ordinary account of two Siamese fighting fish mating is not logically equivalent to the physico-geometrical account of that phenomenon. If offspring never resulted from the series of events we now refer to as 'the mating of Siamese fighting fish', and if, further, offspring did result from some *other* series of events, then the original

series of events, though in no way changed in themselves, would *not* be the mating of Siamese fighting fish. The 'mating' account and the physico-geometrical account are saying different things about what is happening, and hence cannot be logically equivalent.

Similarly, since *every* teleological description is of a phenomenon taking place, as it were, in space-time, it will have a corresponding non-teleological description of the physico-geometrical sort. And these two descriptions will not be logically equivalent because they are saying something different about what is taking place. This point is often put by saying that the teleological account says *more* than the physico-geometrical account. The teleological account says that there is such-and-such a series of events, but then says something else in addition. There is, to my knowledge, no serious dispute about this point. It is generally conceded that 'John went to the store to get some bread' says more than 'John went to the store, and John got some bread'; that 'the shark was pursuing its prey' says more than 'the shark was travelling behind a certain fish at a certain distance'; that 'the missile changed course in order to intercept its target' says more than 'the target changed course and then the missile changed course'—and even that it says more than 'the missile changed course *because* the target changed course'. What is disputed, and what will occupy us here, is what this 'surplus meaning' is, and whether or not it can be expressed in non-teleological terms.

There is some reason to believe that the boundary separating teleological expressions from non-teleological expressions is not a clear, sharp one. Partly for this reason as well as partly for others, there is some dispute about whether certain terms (e.g. 'follow', 'normal', 'important', 'attain') are, always or ever, teleologically pregnant. But this need not concern us initially. There are many expressions which are clearly one or the other, either teleological or not, and since most of the discussion in this area has taken place with due sensitivity to this problem, we can handle the third category, the borderline cases, as they arise.

Most attempts to demonstrate the reducibility of teleological expressions to non-teleological expressions begin with the adoption of a paradigm of teleological phenomena, then produce an abstract schema based on that paradigm and, finally, produce a non-teleological description of that schema. In his recent *Purpose in Nature* [3] John Canfield notes that most of these

[a]ttempts to translate teleology out of biology seem to fall into one of two patterns. Either the system of translation is a *target* schema, or it is a *furnace* schema. Those who construct target schemata take as their guiding example such instances of teleological behavior as a cat chasing a mouse, or, on the mechanistic side, a homing torpedo proceeding to its goal. Furnace schemata, on the

other hand, are modeled after a different kind of teleological behavior, namely, that exemplified by a house equipped with an automatic furnace. This schema best fits such biological behavior as homeostasis of temperature and of blood sugar ([3] p. 5).

Now, it is not clear that these two schemata are mutually exclusive, as Canfield seems to imply. In fact, neither of the first two translation schemata which Canfield presents in *Purpose in Nature* is easily excluded from either category. Nonetheless, the distinction provides two usefully different ways of appreciating translation attempts. Even more useful is Canfield's organisation of the most important attempts to provide a translation schema, and the influence of the structure of *Purpose in Nature* on this paper should be obvious to anyone familiar with that book.

3 REDUCTION THROUGH FEEDBACK

The first attempt we shall consider to provide a rule for translating teleological descriptions into non-teleological ones, which is also the first presented by Canfield, is that offered by Rosenblueth, Wiener and Bigelow in their classic paper [5]. In this paper they offer a translation rule which is most easily seen as the product of a target schema. All teleological behaviour is directed towards a goal and involves 'a continuous feedback from the goal that modifies and guides the behaving object'. Purposeful reactions are those 'which are controlled by the error of the reaction'. Hence, on this interpretation, teleological behaviour 'becomes synonymous with behaviour controlled by negative feedback . . .'. And feedback, being a rather garden variety, causal-mechanical phenomenon is, presumably, completely describable in non-teleological terminology.

There are, however, many difficulties with this characterisation of teleological behaviour, not the least of which is the problem of specifying a goal object without using the teleologically loaded terms 'goal' and 'error'. But the most important difficulty facing this analysis is the one Israel Scheffler calls "The Difficulty of the Missing Goal-Object". ([6] p. 268). Under this heading Scheffler points out that there are many kinds of phenomena which are undeniably teleological but which contain nothing identifiable as a goal-object from which the signals required by the feedback analysis could be emanating.

. . . a man's purpose in groping about in the dark may be to find matches that are not there, his purpose in going to the refrigerator may be to obtain a non-existent apple, he may seek the philosopher's stone, the holy grail, the fountain of youth, or a living pulsing unicorn. In every such case his behavior is clearly purposive and yet in none is this behavior guided by signals emitted from a goal-object . . . ([6] p. 268).

In reply to the possible objection that these examples all involve human intention and explicit decisions, whereas what Rosenblueth, Wiener and Bigelow were trying to analyse was purposive behaviour in non-intended and non-human cases, Scheffler observes that the same problem arises in the cases in which there is no intention, no explicit decision and even no human. Examples of such cases are:

. . . a standing passenger thrusting his foot outwards suddenly in order to keep his balance in a moving train, a rat depressing the lever in his experimental box in order to secure a food pellet, a small infant crying in order to attract mother's attention . . . For we cannot plausibly suppose our train-rider to be receiving guiding signals from some region with which his foot is to be correlated (*sic*). Neither, when the psychologist stops replacing the rat's pellets, can we describe the rat as receiving directive signals from some such pellet. Nor, finally, when mother, expecting baby to sleep, steps out to the corner store is she available for the issuing of signals guiding the infant's behavior toward final correlation with herself ([6] pp. 268-9).

Feedback mechanisms clearly do not provide the desired translation.

4 REDUCTION THROUGH CAUSAL CHAINS

In *Scientific Explanation* [2], R. B. Braithwaite offers a translation rule which, while also best viewed as based on a target schema, avoids the missing-goal-object difficulty of the previous attempt. In place of the feedback device of Rosenblueth *et al.*, Braithwaite employs the notion of the causal chain, nomically determined by initial conditions, to bridge the gulf separating teleology from the more respectable world of ordinary scientific discourse. On this view, the behaviour of a teleological system is viewed as a causal chain connecting an initial event with a final one. The important characteristic of a causal chain for this argument is that

. . . every event in the system is determined by the whole previous state of the system together with the causally relevant factors in the system's environment or field (which will be called field conditions). Then the causal chain *c* of events in [system] *b* throughout a period of time is nomically determined by the initial state *e* of the system together with the totality of field conditions which affect the system with respect to the events in question during the period. Call this set of field conditions *f*. Then for a given system *b* with initial state *e*, *c* is a one-valued function of *f*; that is, for a given *b* and *e*, the causal chain *c* is uniquely determined by *f*—the set of field conditions ([2] p. 330).

Braithwaite takes the defining characteristic of a teleological system to be the plasticity of its behaviour towards a goal, its 'persistence towards the goal under varying conditions'. Hence his task becomes that of representing this plasticity in terms of the undeniably non-teleological concepts of causal chain and field condition. The first, and easy, step is to translate 'the

system is persistent toward the goal under varying conditions' into 'for any given initial state, the system will achieve the goal under a variety¹ of field conditions'. But that leaves the tough problem of eliminating 'achieve' and 'goal' in favour of non-teleological expressions. Braithwaite offers the following:

Now consider the property which a causal chain in a system may possess of ending in an event of type gamma without containing any other event of this type. Call this property the gamma-goal-attaining property, and the class of all causal chains having this property the gamma goal attaining class. Every causal chain which is a member of [the gamma goal attaining class] contains one and only one event of type gamma, and contains this as its final event ([2] p. 330).

So we have: 'for any given initial state of the system there are a variety of field conditions which determine a causal chain in the gamma-goal-attaining class.' This formulation clearly involves no teleologically loaded concepts, and just as clearly requires no signals to be emitted by the goal, and hence avoids the missing-goal-object difficulty.

This formulation does, however, have many difficulties of its own. Five possibly minor ones, which nevertheless deserve mention, are the following. First, there is often a problem in determining which is the terminal event of a causal chain. It is often not clear when it is appropriate to say that a causal chain has ended. Does the causal chain which began when a driver lost control of his car on a wet road end when all of the pieces settle to the ground after the crash, or when the last hubcap stops rolling and all is quiet, or when the driver is resting quietly in the hospital; or is it when he has finally recovered, or when the car has been repaired or, perhaps, both? Or is it only when the driver eventually dies and removes from the causal stream all those neuroses which resulted from the trauma of the accident? It would, I think, be rather arbitrary to choose any of these unqualifiedly as *the* terminal event. Second, it is not obvious that the achievement of a goal must constitute the terminal event (when *that's* clear) of a causal chain. Eating lunch, though clearly an intended and achieved goal, might be just one event in the larger causal chain constituting one's trip to the zoo. Third, there is a problem similar to the first in determining when two events are appropriately of the same type. Clearly the causal chain which produces a shady resting place for an animal can have other shady moments and yet allow that the shady resting place was the goal of the behaviour involved. Hence, shady cannot be the 'type' of event the achievement of the goal is; for if it were, the causal chain would not fit Braithwaite's definition of a gamma-goal-attaining causal chain, because it would then

¹ Strictly, more than one is enough, but the more conditions under which the goal is achieved the more plastic is the behaviour.

contain two events of the same 'type.' The same procedure can be repeated with respect to 'restful' and any other potentially relevant predicate, which makes it difficult to imagine just what 'type' of event finding a shady resting place is in Braithwaite's sense. Fourth, it appears quite likely that non-teleological systems could satisfy the Braithwaite criteria for purposive behaviour if achievement of the goal under as few as two distinct sets of field conditions is allowed as a demonstration of plasticity. For example, a lump of clay would hit the same spot on the floor when dropped from a given place both in still air and in the presence of oppositely directed air currents of the appropriate magnitude and locations. Fifth, it does not seem unreasonable to suggest that there might be occasions on which we would say a system manifested purposive behaviour even though, given its initial state, it was nomically *impossible* for it to achieve the goal to which it aspired. Behaviour might be purposive even though there are *no* possible gamma-goal-attaining causal chains. The fish struggling to get free of its net would seem to be a case of this type. It is struggling toward its goal of freedom even though achievement is impossible merely in virtue of its initial state and the relevant laws. Each of these difficulties is at least sufficient to justify a change in the terms in which the Braithwaite formulation is couched; some might well be more telling than that. None, however, justifies further discussion in this place, largely because of the conclusiveness of the following objection, put forth, again, by Scheffler [6]. This he calls 'The Difficulty of Multiple Goals'.

According to Braithwaite, all that must be done to justify describing the activity of something as teleologically directed towards a certain goal, is to demonstrate the plasticity of the behaviour of that something from the initial state in question toward that goal. And the demonstration of plasticity consists in the specification of at least two sets of field conditions under which the goal is reached from the given initial state. Now, Scheffler's point is merely that, given a system's initial state, there may be sets of field conditions which produce goal gamma, and others which produce goal delta. Accordingly when the system is in this initial state, we should be able to say either (or both) that the system is teleologically directed toward gamma or (and) teleologically directed toward delta. 'This', says Scheffler, 'is exactly what we cannot generally say':

The cat crouching before the vacant mouse-hole is crouching there in order to catch a mouse. Since no mouse is present there will in fact be no goal-attainment. Nevertheless, the cat's behavior is plastic since there are various hypothetical sets of field conditions, each set including one condition positing a mouse within the cat's range such that, in conjunction with the cat's present behaviour, each set determines a mouse-attaining causal chain. On the other hand, there are also

various other hypothetical sets of field-conditions, each set including one positing a bowl of cream within the cat's range, such that, conjoined to the cat's present behavior, each set determines a cream-attaining causal chain. It should therefore be a matter of complete indifference, so far as the present proposal is concerned, whether we describe the cat as crouching before the mouse-hole in order to get some cream. The fact that we reject the latter teleological description while accepting the former is a fact that the present proposal cannot explain ([6] pp. 273-4).

This objection is so telling mostly because it attacks the very foundation of the Braithwaite project. It suggests that *no* analysis will work which takes as its basis a subjunctive statement of goal-attainment (under various possible conditions) of the sort employed in this translation attempt.

5 REDUCTION THROUGH STATE DESCRIPTIONS

Braithwaite contends that his analysis works for furnace-like cases of directly organised behaviour as well as those of the target variety. This presumably would be accomplished by allowing the major variable to be something other than displacement from an object (goal), such as temperature or blood-sugar concentration, and by allowing the plasticity to consist in the maintenance of that variable within a certain range under several sets of field conditions rather than insisting that it tend to zero, as displacement must. Hence, it need not be surprising that the following translation attempt, even though it takes a furnace schema as its model, offers a solution to the very same problems as do the first two models, and is moreover very like Braithwaite's attempt even in some specific points of detail.

This model, which is offered by Ernest Nagel in *The Structure of Science* ([4] p. 410 f.), does, however, avoid many of the difficulties which beset Braithwaite, partly *because* of the difference of his paradigm, and partly through elaborate attention to detail. Because he uses a furnace schema, Nagel is naturally led to express the characteristic feature of teleological systems in a way somewhat different from the previous analyses. He takes this characteristic feature to be the persistence of a system in either manifesting a certain property, condition or mode of behaviour (*G*) or maintaining itself in a state (*G*-state) which causally guarantees that it will manifest that property (etc.) at a future time.¹ But this persistence is with respect to changes *within* the system itself, and explicitly excludes environmental effects on the manifestation of *G*. Hence, the cases of plastic behaviour with respect to environmental variables, which Braithwaite considered, are accommodated by including any causally relevant part of the

¹ The second disjunct of this formulation is, I believe, included explicitly to handle target cases which would otherwise be difficult to accommodate within this analysis.

environment as *part* of the 'directively organized system' in Nagel's terms. So, what is ordinarily considered to *be* the teleological system in these (target) cases, is merely *part* of what Nagel takes to be the 'system' in his analysis.

For this analysis, a system must be analysed in terms of a finite number of non-redundant¹ predicates or state variables which are nomically/physically/causally sufficient to determine whether or not the system manifests *G* or is in a *G*-state.² In these terms Nagel's position is that a system 'will be said to be "directively organised during the interval of time *T* with respect to *G*"' if the system begins that interval in a *G*-state and if, whenever the system undergoes a state-variable change which would *ceteris paribus* take it out of its *G*-state, this induces the other state-variables to change in such a way that the system remains in a *G*-state. A house is a directively organised system on a certain day (*T*) with respect to its temperature if it begins that day in a certain temperature range (i.e. in a *G*-state) and if, whenever (during *T*) the heat-flow out of the house would, if not compensated for, take the house out of the acceptable temperature range, its furnace comes on long enough to effect the required compensation. Nagel adds the condition that the induced state-variable change, in the *absence* of the primary state-variable change, must be one which would take the system out of its *G*-state. The desirability of this condition escapes me, but being part of the analysis it is unreasonable to exclude it.

This analysis does admirably achieve one of its goals: it does not contain any terms that even hint of teleology. It is almost a paradigm of non-teleology. The analysis also clearly avoids the first three of my five 'possibly minor' objections to Braithwaite's model by avoiding talk of terminal events and types of events. But any detailed discussion of how (or whether) Nagel's analysis avoids the rather sophisticated difficulties of the Braithwaite model would be largely wasted here. For in spite of (or, perhaps, because of) his careful attention to detail, Nagel's analysis encounters problems on a very low/coarse/fundamental level. The first and most astonishing is a difficulty so obvious that the previous analyses explicitly avoid it. This is called by Scheffler, in another context,³ 'the difficulty of

¹ This qualification is added merely to avoid formal logical difficulties with the eventual formulation of directive organization.

² Henceforth in this section, 'System *S* either manifests *G* or is in a *G*-state' will be written simply 'System *S* is in a *G*-state'. The first part will be understood unless explicitly noted to the contrary.

³ Scheffler, on first reading, might seem to be imputing this difficulty to Braithwaite's analysis of teleological systems. However, Braithwaite's problem, as Scheffler recognises, is with 'explanation' not 'teleological system'. Hence his analysis of 'teleological system' escapes 'the difficulty of goal-failure' though his analysis of 'teleological explanation' does not.

goal-failure'. The difficulty of goal-failure is most clearly seen by observing that a system can be teleological/goal-directed/directively organised without its *ever* achieving the state towards which it is 'directed'. Scheffler's examples are of a dog pawing at a door while trapped in a place from which it will never escape, and of a cat crouching before an empty mouse-hole. Both cases can be legitimately described teleologically even though the respective goals will never be attained: the dog is pawing in order to get out, the cat crouching in order to catch a mouse. To use the terms of the present analysis: before being consumed, the shark's dinner might have been manifesting behaviour which was clearly directively organised toward the goal of *not* becoming the shark's dinner. Examples can be multiplied indefinitely: trying, which is paradigmatically teleological, does not entail success. On the other hand, before he will allow a system to be called 'directively organized toward *G*', Nagel requires that the manifestation of *G* be causally guaranteed by the system's state description and the relevant laws. I see no easy way for him to escape this difficulty.

The second problem with Nagel's analysis is actually a cluster of closely related methodological difficulties. Nagel leaves unanswered a number of questions which it is his business to answer when treating the problem he claims to be treating. He has given us a formula for telling when a system is 'directively organized during interval of time *T*', but has not said how this relates to important questions couched in less technical terms. For example, when is a system teleological? Only when it is directively organised? Can a teleological system sleep, or does it rather cease to exist for a time? One could offer that the plausible suggestion here is to say that a system is teleological if it is *capable* of directively organised behaviour. But this suggestion raises a host of new problems, and, more importantly, it is only a suggestion. It is not part of Nagel's treatment. And until he provides explicit answers to the important questions about teleology, until he tells us when a system is teleological, when it is legitimate to say 'in order to', when 'goal' is the appropriate description of an object, he has not provided us with anything very useful.

6 THE GENERAL CASE

The above criticism of Nagel, like the criticism of the other models, has consisted in bringing to light rather specific difficulties with the way in which the model goes about producing a translation rule. Little of a general nature has so far been said about the programme of reduction itself. The more general argument has been deferred thus far because a crucial point is more easily developed after looking carefully at some serious attempts to carry out the programme.

The teleological reductionist position is almost always described as claiming that, for every legitimate teleological description of a phenomenon, there can be found a non-teleological description which is logically equivalent to it. Strictly speaking this description is not inaccurate, but it is somewhat misleading. For it seems merely to be maintaining that teleological descriptions can be rewritten, perhaps elliptically, in such a way as to avoid all the traditionally teleological or purposive terms. But upon examination of what the reductionist philosophers are doing, 'non-teleological' might better be replaced by 'un-teleological' or even 'anti-teleological'. These men not only want teleological descriptions to be logically equivalent to something else, they want that something else to be part of the presently accepted conceptual framework of science; they want it to introduce 'nothing new'. This severely restricts what is to count as a 'non-teleological description', and makes the position more interesting and more important, as well as more difficult to defend, than it at first appeared.

It is their desire to avoid anything conceptually novel, to maintain teleology within the domain of *causal* principles, which explains something otherwise rather puzzling done by both Braithwaite and Nagel. The models of teleology offered by these two philosophers might easily have been quite unproblematic behaviourist accounts of teleological systems: Braithwaite's in terms of plastic behaviour with respect to an environment, Nagel's in terms of a system's maintaining a *G*-state in spite of a variety of internal changes. And on the looser, more inclusive interpretation of 'non-teleological' this would have been sufficient. The translation rule would then be: teleological behaviour is behaviour which is plastic with respect to certain environmental obstacles or, alternatively, behaviour in which a system maintains itself in a *G*-state in spite of internal changes normally antithetical to maintaining this *G*-state. But both authors felt compelled to include something in their models which, from the less restrictive view of the translation programme, is completely gratuitous. Braithwaite adds to plasticity the further requirement that 'every event in the [teleological] system is determined by the whole previous state of the system together with the causally relevant factors in the system's environment . . .' ([2] p. 330). Nagel says the same thing in slightly different words when he requires that the 'internal changes', in terms of which his analysis is formulated, be alterations of 'state variables', and that,

whatever the nature of the state variables, in respect to the states they represent *S* is a deterministic system: the states of *S* change in such a way that, if *S* is in the same state at any two different moments, the corresponding states of *S* after equal lapses of time from those moments will also be the same ([4] p. 412).

The requirement that teleology be maintained within the domain of causal principles, that teleology encompass nothing not already within the science of macro-objects, just is the requirement that teleological systems be causally deterministic. All three models presented thus far include this requirement, the first even going so far as to specify the kind of mechanism involved. And just to avoid one possible confusion, I should emphasise here that I think there is a very good chance that all teleological systems that will ever be discovered to exist, will in fact be causally deterministic systems. But to require causal determinism, as part of an analysis of what it is to *be* a teleological system, is to require that a teleological system be causally deterministic by logical necessity. And this dooms the analysis at the outset. For, as Braithwaite and Nagel clearly recognise, what is *essential* to the teleology of a system is the plasticity of its behaviour and its persistence toward a goal.¹ And the plasticity and persistence of a system's behaviour is logically consistent with the causal indeterminacy of that behaviour. When it is said that a system's behaviour is plastic, all that is being asserted is that the system avoids some environmental obstacles. The exact path taken is not important. As long as the goal is reached in a variety of circumstances or, at least, as long as *some* obstacles are circumvented, the behaviour is plastic. The exact path taken to reach the goal, or to avoid the obstacles, can be any one of a rather large number. And if this one were chosen at random, say by a quantum randomiser, the behaviour would be causally indeterminate; it could not be predicted or explained by causal principles. This has the interesting consequence that behaviour can be teleologically determinate, in the sense that it is subsumed under a true teleological generalisation, while at the same time being causally *indeterminate*. A teleological generalisation says only that the behaviour in question is plastic in its persistency toward a goal. No commitment need be made to an exact function relating position, velocity and time.

Perhaps this is what Morton Beckner recognised when he stopped in the middle of a feedback-mechanistic analysis of teleology to say:

It should be clearly recognized that we have here an empirical relation between purposive behavior and teleological systems, and not a logical relation between the concepts 'purposive behavior' and 'teleological system'. ([I] p. 144).

This is a very puzzling thing for him to say, and appears to make nonsense of his analysis. For it would seem that teleological systems just are, by definition if you will, those things which behave purposively. This is how they are recognised, distinguished. But Beckner probably saw that some such disavowal follows automatically from any attempt to give a logical

¹ This is its persistence in a *G*-state for Nagel.

analysis of teleological systems in causal-mechanistic terms. So he included this statement in his treatment, perhaps hoping that public admission would go most of the way toward solving the problem; which, of course, it does not.

All of the scholars mentioned in this paper make a point of saying that it is sometimes not only legitimate but *useful* to think of a subject matter in teleological terms. Most of them justify their excursion into this area in this way. But, and this is the kernel of the present essay, if it is ever useful to think of something as teleological, then it is useful to think of it as something which is not necessarily causally deterministic. A teleological expression is not just an abbreviation for something complex but essentially causal. Even if all teleological or purposive behaviour is in fact the product of a causal mechanism, and hence is *explainable in terms of* causal principles, it does not follow that the teleological description of such behaviour is *reducible to* a causal description. Teleological expressions are not 'translatable into' or 'saying the same thing as' some causal expression. The two do not 'have the same cognitive content', and they differ in more than mere 'emphasis'. If teleological expressions are *ever* scientifically useful, they are so because of the way teleological concepts organise an area of scientific interest, not because of their equivalence to causal expressions.

Rochester Institute of Technology
Rochester, N. Y.

REFERENCES

1. Beckner, M. *The Biological Way of Thought* (Columbia University Press, 1959).
2. Braithwaite, R. B. *Scientific Explanation* (Cambridge University Press, 1953).
3. Canfield, J. *Purpose in Nature* (Prentice-Hall, 1966).
4. Nagel, E. *The Structure of Science* (Harcourt, Brace and World, 1961).
5. Rosenblueth, Wiener and Bigelow, 'Behavior, Purpose and Teleology'. *Phil. Sci.* 10 (1943), 18-24.
6. Scheffler, I. 'Thoughts on Teleology', *Brit. J. Phil. Sci.* 9 (1959), 265-84.

Some Aspects of Carnap's Theory of Inductive Inference

by CARL-ERIK SÄRNDAL

I INDUCTIVE LOGIC AND STATISTICAL INFERENCE

In this paper we shall be concerned with certain problems of inductive logic that are of importance in theories of statistical inference. Following Carnap (1950), we shall deal with a logical setup in the form of a language where certain sentences of evidence e and sentences of hypothesis h are considered. There is assumed to exist a logical relationship between e and h , which is numerically measurable through the quantity $c(h, e)$, called the degree of confirmation of h on e . This measure is such that $0 \leq c(h, e) \leq 1$.

In statistical theory, two examples of inductive logic are of special interest, namely, those of direct and indirect inference. In direct inference, e is a statement that relates to the population, while h relates to a sample drawn from the population. In indirect inference, on the other hand, the situation is reversed so that e is a statement concerning the sample, while h relates to the population. In the cases we shall consider, the sentences e and h will usually deal with frequencies of certain properties of the objects that make up the population or the sample.

A theory of statistical inference may be said to be a collection of methods for drawing conclusions about the population on the basis of a sample, usually drawn at random from the population. Judging from the above, it appears that the logic of indirect inference is immediately tailored to suit a theory of statistical inference. Nevertheless, the dominating statistical inference theory of today, that of Neyman-Pearson-Wald, is founded on the logic of direct inference.

Some of the terminology of Carnap (1950, 1952, 1959) will be adopted in this paper. Following Carnap (1950, pp. 8-11) we shall distinguish between *classificatory*, *comparative*, and *quantitative* concepts. As an example of each of these three categories, take the concepts 'warm', 'warmer', and 'temperature', respectively. A concept such as 'warm', or its negation 'not warm', can be used to classify each object of any collection of objects. We shall say that 'warm' is a *classificatory property*. We assume that through the concept of 'temperature' a certain standard or norm for

'warm' has been defined. This may be given, for example, as an interval on the centigrade scale. Hence, the property 'warm' can be tagged on to any object whose measured temperature is found to comply with the standard.

Consider the classificatory properties 'cold', 'luke-warm', 'warm', 'hot', all of which are assumed to be defined through the concept 'temperature' in such a way that each object must possess one and only one of these four properties. Such a set of properties is called a *family of related properties* (Carnap, 1950, p. 76). Clearly, the properties 'small', 'wooden', and 'spherical', for example, do not form a family of related properties. Unlike Carnap, whose premises are more general, we shall limit consideration to such properties that form a family of related properties with reference to a certain quantitative concept. As in the example of 'cold', 'luke-warm', 'warm', and 'hot', we assume that the properties in the family can be arranged in a spectrum, where each property corresponds to a higher range of the scale of measurement for the underlying quantitative concept than the preceding one. We shall also use the term *pairwise exclusive and exhaustive* for a set of properties such that an object must possess one and only one property in the set.

Consider now the comparative concept 'warmer', which is always a dyadic relationship. It bears connection to the quantitative concept 'temperature' in such a way that an object *a* is said to be 'warmer' than an object *b* if the temperature of *a* exceeds that of *b*. In a collection of more than two objects, we assume the possibility of performing all pairwise comparisons with reference to the concept 'warmer'. Thus, for any collection of objects, we assume the existence of a ranking of the objects by means of a quantitative concept. Furthermore, we shall assume that no two objects are exactly alike as measured by the quantitative concept. If in a collection of objects, an object *a* is found to be 'warmer' compared to *m* of the objects but 'not warmer' compared to the rest, we shall say that *a* possesses the *order property* of ranking as number *m*. The order properties are pairwise exclusive and exhaustive, and, moreover, each order property is possessed by one and only one object.

Two critical components in the argument used by Carnap in arriving at his theory of inductive inference are the nature of the logical measure function and the choice of the primitive properties. As far as the logical measure function is concerned, Carnap's theory gives equal measure to each structure description. Furthermore, all objects are treated on a par, and equal measure is given to each isomorphic state description within a structure description (for details, see Tintner, 1949, and Carnap, 1950).

Another central issue in Carnap's theory is the choice of the primitive properties. The possibility of defining primitive properties in a number of

ways introduces a degree of arbitrariness into Carnap's theory of inductive inference. The consequences of this arbitrariness for the particular situations we shall deal with will be discussed later in the paper.

Assume that a set of primitive properties (classificatory or order properties) $P_1, P_2, \dots, P_m$ has been agreed on. They are assumed to form a family of related properties, i.e. they are pairwise exclusive and exhaustive. A new family of related properties $Q_1, Q_2, \dots, Q_k$ ($k < m$) may then be formed by defining each Q_i as the disjunction of a given number ω_i ($\omega_i \geq 1$) of the primitive properties $P_1, P_2, \dots, P_m$. Each P_i is present in one and only one of the Q_i . Then $Q_1, Q_2, \dots, Q_k$ are called *molecular properties*, and the positive integer ω_i is called the *logical width* of Q_i . We then have $\sum_{i=1}^k \omega_i = m$.

In the language that we shall be concerned with, a typical sentence will state the cardinal number for the frequency of a certain molecular or primitive property. Such a sentence will be written symbolically in the form of a property symbol preceded by the letter 'N' (for population frequency) or 'n' (for sample frequency) and followed by the sign '=' and an integer representing the population or sample frequency. For example, the sentence ' $NP_i = m$ ' states that the number of objects in the population with the property P_i is equal to the integer m , and ' $nQ_i = s$ ' states that the number of objects in the sample with the property Q_i is equal to s . The conjunction of such sentences for each of a number of pairwise exclusive and exhaustive properties forms a *statistical distribution* (Carnap, 1950, p. 111; 1952, p. 38). The statements of evidence e and hypothesis h that we shall consider will usually be in the form of statistical distributions.

Individual objects are denoted by ' a_i ' ($i = 1, 2, \dots$). Sentences describing individual objects will not be needed in this paper. Logical connectives between two or more properties or sentences will be used in the following way. The sign of disjunction will be written 'V' with the meaning of 'or' (in its non-exclusive sense). The disjunction $A_1 \vee A_2 \vee \dots \vee A_k$ will be written

$$\vee A_i \ (i = 1, 2, \dots, k).$$

The sign of conjunction will be written with the meaning of 'and'. The conjunction $A_1 \cdot A_2 \cdot \dots \cdot A_k$ will be written

$$\cdot A_i \ (i = 1, 2, \dots, k).$$

2 CARNAP'S SYSTEM FOR DIRECT AND INDIRECT INFERENCE

The Carnapean system for direct and indirect inference was developed by Carnap (1950) and was expounded to statisticians by Tintner (1949). In this section we shall review the results of Carnap's theory of inference for

classificatory properties in the direct and indirect cases. In Sections 3, 4, and 5 we shall turn to an application of Carnap's system of inference in the case of order properties. *

The population is assumed to consist of N objects $a_1, a_2, \dots, a_N$. A sample of n objects is assumed to be drawn from the population without replacement. Let there be a family of related properties $P_1, P_2, \dots, P_\kappa$ defined with reference to a certain quantitative concept. We choose to take each of the P_i as a primitive property. Consider k molecular properties R_i ($i = 1, 2, \dots, k$), each formed by disjunction of a number of P_i . Let the logical width of R_i be ω_i , where $\omega_i \geq 1$, and assume that the properties $R_1, R_2, \dots, R_k$ form a new family of related properties. Thus $\sum_{i=1}^k \omega_i = \kappa$.

In the case of direct inference, consider the sentences

$$\begin{aligned} e: (NR_i = s_i) \quad (i = 1, 2, \dots, k), \\ h: (nR_i = r_i) \quad (i = 1, 2, \dots, k). \end{aligned}$$

Then Carnap's result for the degree of confirmation of h on e is (see Carnap, 1950, p. 495 and Tintner, 1949)

$$c(h, e) = \frac{m(e \cdot h)}{m(e)} = \frac{\prod_{i=1}^k \binom{s_i}{r_i}}{\binom{N}{n}} \quad (2.1)$$

where $m(e)$ and $m(e \cdot h)$ are the logical measures of the sentences e and $e \cdot h$, respectively, and

$$\begin{aligned} m(e) &= \prod_{i=1}^k \binom{s_i + \omega_i - 1}{\omega_i - 1} / \binom{N + \kappa - 1}{\kappa - 1}, \\ m(e \cdot h) &= \prod_{i=1}^k \binom{s_i + \omega_i - 1}{\omega_i - 1} \binom{s_i}{r_i} / \binom{N + \kappa - 1}{\kappa - 1} \binom{N}{n}. \end{aligned} \quad (2.2)$$

Formula (2.1) does not depend on the logical widths ω_i . We must have $s_i \geq r_i \geq 0$ for each i . The relations between the s_i and N on the one hand, and between r_i and n on the other, are expressed by the equations

$$\begin{aligned} \sum_{i=1}^k s_i &= N, \\ \sum_{i=1}^k r_i &= n. \end{aligned} \quad (2.3)$$

In the case of indirect inference, consider the sentences

$$\begin{aligned} e: (nR_i = r_i) \quad (i = 1, 2, \dots, k), \\ h: (NR_i = s_i) \quad (i = 1, 2, \dots, k). \end{aligned}$$

According to Carnap (1950, p. 570), and Tintner (1949), the degree of confirmation of h on e is given by

$$c(h, e) = \frac{m(e \cdot h)}{m(e)} = \frac{\prod_{i=1}^k \binom{s_i + \omega_i - 1}{r_i + \omega_i - 1}}{\binom{N + \kappa - 1}{n + \kappa - 1}}, \quad (2.4)$$

where $s_i \geq r_i \geq 0$ for each i , and (2.3) holds. The logical measure of e is

$$m(e) = \prod_{i=1}^k \binom{r_i + \omega_i - 1}{\omega_i - 1} / \binom{n + \kappa - 1}{\kappa - 1},$$

and $m(e \cdot h)$ is given by (2.2), which has the alternative form

$$m(e \cdot h) = \prod_{i=1}^k \binom{r_i + \omega_i - 1}{\omega_i - 1} \binom{s_i + \omega_i - 1}{r_i + \omega_i - 1} / \binom{n + \kappa - 1}{\kappa - 1} \binom{N + \kappa - 1}{n + \kappa - 1}.$$

Unlike the direct formula (2.1), the indirect formula (2.4) depends on the logical setup through the logical widths ω_i . To connect the numbers $s_i + \omega_i - 1$ with $N + \kappa - 1$ on the one hand and $r_i + \omega_i - 1$ with $n + \kappa - 1$ on the other, we have the rules of thumb

$$\begin{aligned} \sum_{i=1}^k (s_i + \omega_i - 1) + (k - 1) &= N + \kappa - 1, \\ \sum_{i=1}^k (r_i + \omega_i - 1) + (k - 1) &= n + \kappa - 1. \end{aligned} \quad (2.5)$$

Note that the number of properties less one has to be added in the left member in order to arrive at the integers $N + \kappa - 1$ and $n + \kappa - 1$.

3 ORDER RELATIONS IN THE POPULATION AND IN THE SAMPLE

Consider first the N objects $a_1, a_2, \dots, a_N$ that constitute the population. We assume the possibility of making comparisons among all pairs of objects with respect to a certain quantitative concept (like, for example, 'temperature'). As a result of these comparisons we may assume the existence of an arrangement of the objects according to ascending values of the quantitative concept. The arrangement will be called an *ordered population sequence*. This sequence may be written as

$$a_{T_1} a_{T_2} a_{T_3} \dots a_{T_N}, \quad (3.1)$$

where $T_1, T_2, \dots, T_N$ is a permutation of the integers $1, 2, \dots, N$. (In the example of 'temperature', the implication is that a_{T_i} is 'warmer' than $a_{T_{i-1}}$ for $i = 2, 3, \dots, N$). By the i th *population object* or the *object of population order i* we shall mean the i th object from the left extreme of the ordered population sequence, i.e. the object a_{T_i} . (When there is no risk of an ambiguous interpretation, the word 'population' may be deleted in the terms just defined, and in such cases we shall speak simply of 'the object of order i ', for example.) The property of being of population order i will be denoted by P_i^* and will be referred to as a *population order property*.

We have assumed and shall continue to assume that for every collection of objects, the arrangement into an ordered sequence is possible without any 'ties'. In other words, it is assumed impossible for any two objects to be exactly alike with respect to the quantitative concept that forms the

basis for the ordering. As a consequence, for each $i = 1, 2, \dots, N$, there will be one and only one object having the property P_i^* .

Consider next a sample from the population, i.e. a subset of n objects chosen from the set of objects $\{a_1, a_2, \dots, a_N\}$. Assume the possibility of making comparisons among all pairs of sample objects. Thus we can assume the existence of an arrangement of the sample objects according to ascending values of the quantitative concept. The arrangement will be referred to as the *ordered sample sequence*. This will be written as

$$a_{i_1} a_{i_2} a_{i_3} \dots a_{i_n}. \quad (3.2)$$

By the i th *sample object* or the *object of sample order i* we shall mean the i th object from the left extreme of the ordered sample sequence. (In these terms, the word 'sample' may be deleted if there is no risk of ambiguity.) The property of being of sample order i will be denoted by ' p_i^* '. We shall call p_i^* a *sample order property*.

4 DIRECT INFERENCE FOR ORDER PROPERTIES

We shall devise a framework for direct inference for order properties. In order to do this, a definition of primitive property is first needed. We shall choose to take each of the population order properties P_i^* ($i = 1, 2, \dots, N$) as a primitive property.

For any set of k integers $c_1, c_2, \dots, c_k$ ($1 \leq k \leq n$) such that

$$0 = c_0 < c_1 < c_2 < \dots < c_k < c_{k+1} = N+1, \quad (4.1)$$

consider the primitive properties $P_{c_i}^*$ ($i = 1, 2, \dots, k$). Define the disjunct sets of integers

$$M_i = \{c_{i-1}+1, c_{i-1}+2, \dots, c_i-1\} \quad (i = 1, 2, \dots, k+1),$$

and let

$$\gamma_i = c_i - c_{i-1} \quad (i = 1, 2, \dots, k+1).$$

The set M_i is empty if $\gamma_i = 1$. For each non-empty set M_i (i.e. a set for which $\gamma_i > 1$), define for $i = 1, 2, \dots, k+1$ the molecular properties

$$R_i^* = \vee P_v^* \quad (v \in M_i).$$

The logical width of R_i^* is $\gamma_i - 1$.

Consider further the sample order properties $p_{b_i}^*$ for integers b_i ($i = 1, 2, \dots, k$) such that

$$0 = b_0 < b_1 < b_2 < \dots < b_k < b_{k+1} = n+1, \quad (4.2)$$

and set

$$\beta_i = b_i - b_{i-1} \quad (i = 1, 2, \dots, k+1).$$

Consider the sentence stating that a certain set of k objects are of population orders $c_1, c_2, \dots, c_k$, and that *the same* objects are included in the sample and, moreover, that they are of sample orders $b_1, b_2, \dots, b_k$, respectively.

This sentence is equivalent to stating that the object of population order c_i is identical to the object of sample order b_i for $i = 1, 2, \dots, k$, which we write symbolically as

$$(a_{T_{c_i}} \equiv a_{b_i}) \quad (i = 1, 2, \dots, k), \quad (4.3)$$

using the notation of the sequences (3.1) and (3.2). Choosing to treat this situation as one of direct inference, the sentence (4.3) may be expressed as the conjunction of the sentences e and h , where

$$\begin{aligned} e: & (NR_i^* = \gamma_i - 1) \quad (i = 1, 2, \dots, k+1) \quad \cdot (NP_{c_i}^* = 1) \quad (i = 1, 2, \dots, k), \\ h: & (nR_i^* = \beta_i - 1) \quad (i = 1, 2, \dots, k+1) \quad \cdot (nP_{c_i}^* = 1) \quad (i = 1, 2, \dots, k). \end{aligned} \quad (4.4)$$

Note that we may have $\gamma_i = 1$ for certain i so that R_i^* is not defined. It is practical and legitimate, still, to use the formulation (4.4), in which we make the convention that the property R_i^* should be regarded as non-existent for each case where $\gamma_i = 1$. If $\gamma_i = 1$, then the population and sample frequencies are, of course, both equal to zero.

We shall employ the Carnapean formula for direct inference in a form suited to these order properties. There is only one structure description in this situation, since each of the N primitive properties is possessed by one and only one of the N population objects. The number of state descriptions within this structure description is $N!$. In (4.4) we are dealing with $2k+1$ molecular properties (at most), namely, R_i^* ($i = 1, 2, \dots, k+1$) and $P_{c_i}^*$ ($i = 1, 2, \dots, k$). The logical width is equal to $\gamma_i - 1$ for R_i^* ($i = 1, 2, \dots, k+1$) and unity for $P_{c_i}^*$ ($i = 1, 2, \dots, k$), but, as we noted in Section 2, in direct inference logical widths play no visual role in the formula for degree of confirmation. We obtain therefore, in analogy with (2.1), the degree of confirmation as

$$c(h, e) = \frac{\prod_{i=1}^{k+1} (\gamma_i - 1)}{\prod_{i=1}^k (\beta_i - 1)} \frac{\prod_{i=1}^k (1)}{\binom{N}{n}} = \frac{\prod_{i=1}^{k+1} (\gamma_i - 1)}{\prod_{i=1}^k (\beta_i - 1)} \frac{1}{\binom{N}{n}}, \quad (4.5)$$

where

$$\beta_i \geq 1 \quad (i = 1, 2, \dots, k+1), \quad \sum_{i=1}^{k+1} \beta_i = n+1, \quad (4.6)$$

$$\gamma_i \geq 1 \quad (i = 1, 2, \dots, k+1), \quad \sum_{i=1}^{k+1} \gamma_i = N+1, \quad (4.7)$$

$$\gamma_i \geq \beta_i \quad (i = 1, 2, \dots, k+1). \quad (4.8)$$

In correspondence with (2.3) we have in (4.5) the relations

$$\sum_{i=1}^{k+1} (\gamma_i - 1) + \sum_{i=1}^k 1 = N, \quad \sum_{i=1}^{k+1} (\beta_i - 1) + \sum_{i=1}^k 1 = n.$$

If the degree of confirmation is summed for every hypothesis h for

which (4.6) and (4.8) hold, for a set of fixed γ_i satisfying (4.7), then we obtain

$$c(h', e') = \sum c(h, e) = \binom{N-k}{n-k} / \binom{N}{n}, \quad (4.9)$$

where

$$\begin{aligned} e' &: (NP_{c_i}^* = 1) \quad (i = 1, 2, \dots, k), \\ h' &: (nP_{c_i}^* = 1) \quad (i = 1, 2, \dots, k). \end{aligned}$$

Then $e' \cdot h'$ states simply that the objects of population order c_i ($i = 1, 2, \dots, k$) are included in the sample, and the degree of confirmation of h' on e' is given by (4.9).

For every term of the summation in (4.9), the sentence e of (4.4) is the same, since the γ_i are fixed. In the summation (4.9) we take into account all possible hypotheses h that can be sustained to a positive degree of confirmation by the fixed evidence e with properties R_i^* (of logical width $\gamma_i - 1$) and $P_{c_i}^*$ (of logical width 1).

If, on the other hand, $c(h, e)$ given by (4.5) is summed for all possible sentences of evidence e and hypotheses h for which (4.7) and (4.8) hold, for a set of fixed β_i satisfying (4.6), we obtain

$$\sum c(h, e) = 1.$$

The summing to unity is characteristic of probability distributions, and (4.5) does represent a probability distribution of fundamental interest in one field of direct statistical inference, namely, the theory of order statistics. It should be observed, however, that when the γ_i run through all possible values of the summation, both of the sentences e and h change in (4.4). The properties R_i^* and $P_{c_i}^*$ are different for each new set of γ_i -values.

5 INDIRECT INFERENCE FOR ORDER PROPERTIES

We shall create a framework for indirect inference for order properties to which Carnap's formula for indirect inference will be applied. In order to do this, a definition of 'primitive property' will first be given in reference to this particular situation. This definition will differ from the one used in Section 4 for the direct case.

We assume as before the existence of an ordered sample sequence (3.2) and an ordered population sequence (3.1). Each sample object is identical to one of the population objects. Let the set of population orders corresponding to the sample objects be $m_1, m_2, \dots, m_n$, a set of integers such that

$$0 = m_0 < m_1 < m_2 < \dots < m_n < m_{n+1} = N + 1.$$

The objects of population order ν such that

$$m_{i-1} < \nu < m_i \quad (5.1)$$

define what will be called *the i th block*. The collection of all objects such that (5.1) holds for some i , $i = 1, 2, \dots, n+1$, comprises all the objects *not* included in the sample. There are $N-n^*$ such objects. The property of belonging to the i th block is chosen as a primitive property, and it will be denoted by ' ϕ_i^* '.

We define the molecular properties

$$q_i^* = \forall \phi_v^* (v \in M_i),$$

where

$$M_i = \{b_{i-1}+1, b_{i-1}+2, \dots, b_i\} (i = 1, 2, \dots, k+1),$$

and $b_0, b_1, \dots, b_k, b_{k+1}$ are integers satisfying the order relation (4.2). Let

$$\beta_i = b_i - b_{i-1} (i = 1, 2, \dots, k+1).$$

The logical width of q_i^* is β_i . The β_i satisfy the relations (4.6). Note that the sentences

$$nq_i^* = 0, n\phi_i^* = 0$$

must always be true for every q_i^* and every ϕ_i^* , since there is no sample object having any of the properties q_i^* .

Let $c_0, c_1, \dots, c_k, c_{k+1}$ be integers such that (4.1) holds. Set

$$\gamma_i = c_i - c_{i-1} (i = 1, 2, \dots, k+1).$$

The γ_i then fulfil (4.7), and we shall assume that (4.8) is also satisfied.

Consider again the sentences stating that a certain set of k objects are of sample orders $b_1, b_2, \dots, b_k$, and that *the same* set of objects are of population orders $c_1, c_2, \dots, c_k$, respectively. Symbolically, these identities can be expressed by (4.3). We shall now treat the situation as one of indirect inference. The sentence (4.3) is equivalent to the conjunction of e and h , where

$$\begin{aligned} e: & (nq_i^* = 0) (i = 1, 2, \dots, k+1), \\ h: & (Nq_i^* = \gamma_i - \beta_i) (i = 1, 2, \dots, k+1). \end{aligned} \quad (5.2)$$

In order to obtain the degree of confirmation of h on e , we shall apply Carnap's formula for indirect inference in the case of $k+1$ properties. In analogy with (2.4) we obtain

$$c(h, e) = \prod_{i=1}^{k+1} \left(\frac{\gamma_i - \beta_i + \beta_i - 1}{0 + \beta_i - 1} \right) / \binom{N}{n} = \prod_{i=1}^{k+1} \left(\frac{\gamma_i - 1}{\beta_i - 1} \right) / \binom{N}{n}, \quad (5.3)$$

where the β_i and the γ_i are restricted by (4.6), (4.7), and (4.8). In analogy with the relations (2.5) we have in (5.3)

$$\begin{aligned} \sum_{i=1}^{k+1} (\gamma_i - \beta_i + \beta_i - 1) + k &= N, \\ \sum_{i=1}^{k+1} (0 + \beta_i - 1) + k &= n. \end{aligned}$$

The result (5.3) is identically the same as the direct formula (4.5). This was achieved by using suitable and differing definitions of a primitive property in the direct and indirect cases.

Consider all hypotheses h given by (5.2) for a set of fixed β_i satisfying (4.6). If the degree of confirmation $c(h, e)$ given by (5.3) is summed for all γ satisfying (4.7) and (4.8), then we obtain

$$\sum c(h, e) = 1,$$

a characteristic feature of a probability distribution. For every term of this summation, the sentence e of (5.2) is the same, and the properties q_i^* (of logical width β_i) dealt with in e are fixed. The summation therefore implies the addition of the degree of confirmation of all possible hypotheses h that can be sustained by the fixed evidence e .

Let us compare the interpretation of the same summation given at the end of Section 4. There we remarked on the interest of (5.3) (or the identical formula (4.5)) as a probability distribution in the theory of order statistics. This theory, however, is as a form of direct statistical inference, i.e. there is nothing of a 'Bayesian' nature about it. This is somewhat surprising since, in light of the argument of this section, the probability distribution (5.3) (or (4.5)) seems to lend itself naturally to an indirect interpretation, inasmuch as one makes inference about the population orders of a set of objects on the basis of certain stated facts about the sample, expressed in terms of the fixed sample orders β_i .

At first sight it appears convenient to identify a probability distribution as the set of values $c(h, e)$ that is obtained if one considers all possible hypotheses h that can be sustained to a positive degree of confirmation by a certain fixed sentence of evidence e . It is clear that this narrow view of a probability distribution should not be accepted, for then the probability distributions occurring in the theory of order statistics would be indirect rather than direct.

6 INFERENCE SYSTEMS FOR TWO CLASSIFICATORY PROPERTIES

In this section we shall see how the inference systems for classificatory and order properties are interrelated in the case of two classificatory properties. Consider the pairwise exclusive and exhaustive classificatory properties $P_1, P_2, \dots, P_\kappa$, and define the properties

$$Q_\omega = \vee P_i \ (i = 1, 2, \dots, \omega) \quad (6.1)$$

and

$$\bar{Q}_\omega = \vee P_i \ (i = \omega + 1, \omega + 2, \dots, \kappa), \quad (6.2)$$

where the integer ω is such that $1 \leq \omega \leq \kappa - 1$. The logical widths of Q_ω and $\bar{Q}_\omega$ are ω and $\bar{\omega} = \kappa - \omega$, respectively.

We assume that the primitive properties $P_1, P_2, \dots, P_\kappa$ form a family of related properties arranged in a spectrum so that P_i corresponds to a higher range of scale of the underlying quantitative concept than does P_{i-1} . In other words, recalling the ordered population sequence (3.1), we assume that the objects having the property P_1 would be found at the left extreme of the sequence, then, moving towards the right, the objects having property P_2 follow, and so on. The objects having the property P_κ would be found at the right extreme of the sequence. Thus, if Q_ω is defined by (6.1) we have that the objects with the classificatory property Q_ω are identically the same as those with order property

$$Q_{d_\nu}^* = \vee P_i^* \quad (i \leq d_\nu)$$

for $\nu = 1, 2, \dots, \kappa-1$ and some set of integers $d_1, d_2, \dots, d_{\kappa-1}$ such that

$$0 \leq d_1 \leq d_2 \leq \dots \leq d_{\kappa-1} \leq N.$$

It follows, of course, that in the ordered sample sequence (3.2), counting from the left, we would also have the objects with property P_1 at the left extreme, then the objects with property P_2 would follow, and so on.

Consider the properties Q_ω and $\bar{Q}_\omega$, defined by (6.1) and (6.2), respectively, for some fixed ω , $1 \leq \omega \leq \kappa-1$. For simplicity, write ' Q ' for Q_ω and ' $\bar{Q}$ ' for $\bar{Q}_\omega$. Then in the ordered sequences (3.1) and (3.2) we would have, counting from the left, a number of objects with property Q (that is, if there are any objects with this property at all), and the remaining objects, located toward the right flank of the sequence, would have property $\bar{Q}$.

For direct inference with regard to the classificatory properties Q and $\bar{Q}$ consider the sentences

$$e: (NQ = \gamma) \cdot (N\bar{Q} = \bar{\gamma}),$$

$$h: (nQ = \beta) \cdot (n\bar{Q} = \bar{\beta}),$$

where

$$\gamma + \bar{\gamma} = N, \beta + \bar{\beta} = n.$$

Carnap's formula (2.1) gives

$$c(h, e) = \binom{\gamma}{\beta} \binom{\bar{\gamma}}{\bar{\beta}} / \binom{N}{n} \quad (6.3)$$

For a fixed value of γ , the possible values of β in (6.3) are

$$\beta = u_0, u_0 + 1, \dots, u_1 \quad (6.4)$$

where $u_0 \leq u_1$ and $u_0 = \max(0, n - N + \gamma)$, $u_1 = \min(\gamma, n)$.

For indirect inference with respect to Q and $\bar{Q}$, consider the sentences

$$e: (nQ = \beta) \cdot (n\bar{Q} = \bar{\beta}),$$

$$h: (NQ = \gamma) \cdot (N\bar{Q} = \bar{\gamma}).$$

By Carnap's formula (2.4) we have

$$c(h, e) = \binom{\gamma + \omega - 1}{\beta + \omega - 1} \binom{\bar{\gamma} + \bar{\omega} - 1}{\bar{\beta} + \bar{\omega} - 1} / \binom{N + \kappa - 1}{n + \kappa - 1} \quad (6.5)$$

where ω and $\bar{\omega}$ are the logical widths of Q and $\bar{Q}$, respectively, and $\omega + \bar{\omega} = \kappa$. For a fixed value of β , the possible values of γ in (6.5) are

$$\gamma = \beta, \beta+1, \dots, N-n+\beta. \quad (6.6)$$

In the following we shall also be dealing with sentences involving order properties, which will be particular cases of the general sentence (4.3). In this and the following sections we shall interpret inference for order properties in the direct sense according to Section 4. (We note in passing that the same degree of confirmation would be obtained by the indirect logic according to Section 5.)

The objects of the population or the sample can be described in terms of classificatory properties as well as in terms of order properties. As a consequence, certain sentences stated in terms of classificatory properties may be found to be logically equivalent to some other sentences expressed in terms of order properties. As an illustration, we observe how the result (6.3) may be obtained by a line of reasoning that involves order properties. Consider the sentences

$$\begin{aligned} e: (NQ = \gamma) \cdot (N\bar{Q} = \bar{\gamma}), \\ h: (nQ = \beta) \cdot (n\bar{Q} = \bar{\beta}). \end{aligned}$$

The conjunction $e \cdot h$ is equivalent to the conjunction of all sentences

$$(a_{T_\nu} \equiv a_{t_\beta}) \cdot (a_{T_\mu} \equiv a_{t_{\beta+1}}) \quad (6.7)$$

such that the integers ν and μ satisfy

$$\beta \leq \nu \leq \gamma < \mu \leq N - \bar{\beta} + 1. \quad (6.8)$$

We have by direct inference according to (4.5) that the degree of confirmation of (6.7) is given by

$$\binom{\nu-1}{\beta-1} \binom{N-\mu}{n-\beta-1} / \binom{N}{n}$$

for ν and μ fulfilling (6.8). Summing over all ν and μ satisfying (6.8) for fixed β and γ , we obtain

$$c(h, e) = \sum_{\nu, \mu} \binom{\nu-1}{\beta-1} \binom{n-\mu}{n-\beta-1} / \binom{N}{n} = \binom{\gamma}{\beta} \binom{\bar{\gamma}}{\bar{\beta}} \binom{N}{n}$$

The result confirms the direct inference formula (6.3), as the case should be, since the statements are equivalent.

Next, we attempt to make indirect inference for the classificatory properties of Q and $\bar{Q}$ on basis of the results obtained in the case of direct inference for order properties. Consider the evidence

$$e: (nQ = \beta) \cdot (n\bar{Q} = \bar{\beta}). \quad (6.9)$$

Further, consider the sentence

$$a_{T_\nu} \equiv a_{t_\beta},$$

for which the degree of confirmation is obtained from (4.5) by direct logic as

$$\binom{\nu-1}{\beta-1} \binom{N-\mu}{n-\beta} / \binom{N}{n}, \quad (6.10)$$

and the sentence

$$a_{T_\mu} \equiv a_{t_{\beta+1}}$$

with the degree of confirmation obtained from (4.5) as

$$\binom{\mu-1}{\beta} \binom{N-\mu}{n-\beta-1} / \binom{N}{n}. \quad (6.11)$$

Concerning NQ , the population number with the property Q , we conclude from e that it must satisfy

$$\nu \leq NQ < \mu,$$

where ν is the population order of the last Q of the ordered sample sequence and μ is the population order of the first $\bar{Q}$ of the ordered sample sequence. Consider the sentence

$$h: (NQ = \gamma) \cdot (N\bar{Q} = \bar{\gamma}).$$

On evidence somewhat stronger than as stated by (6.9), we would actually be in a position to solve the problem of indirect inference for classificatory properties simply by reference to the case of direct inference for order properties. For if, on the one extreme, it were true that the object a_{t_β} is identical to the last Q (counting from the left, as always) of the ordered population sequence, then $\gamma = \nu$, and the degree of confirmation of h on e would be given by (6.10) for $\nu = \gamma$, i.e.

$$c(h, e) = \binom{\gamma-1}{\beta-1} \binom{\bar{\gamma}}{\bar{\beta}} / \binom{N}{n} \quad (6.12)$$

If, at the other extreme, it were true that the object $a_{t_{\beta+1}}$ is identical to the first $\bar{Q}$ of the ordered population sequence, then $\gamma = \mu - 1$, and (6.11) would give the degree of confirmation of h on e as

$$c(h, e) = \binom{\gamma}{\beta} \binom{\bar{\gamma}-1}{\bar{\beta}-1} / \binom{N}{n} \quad (6.13)$$

Generally, neither $\gamma = \nu$ nor $\gamma = \mu - 1$ is the case, but rather the truth lies somewhere in between. For future reference, we form the expression

$$c(h, e) = \binom{\gamma+\delta-1}{\beta+\delta-1} \binom{\bar{\gamma}+\bar{\delta}-1}{\bar{\beta}+\bar{\delta}-1} / \binom{N}{n}, \quad (6.14)$$

where $\delta = 0$ or 1 , and $\delta + \bar{\delta} = 1$. It is easily seen that (6.12) obtains for $\delta = 0$ and that (6.13) obtains for $\delta = 1$.

7 RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

In this and the following sections we shall discuss the probability distributions associated with certain random variables defined in connection to

some of the formulas for degree of confirmation obtained in earlier sections. We shall speak about a probability distribution as being of the direct type or the indirect type according as the corresponding inference argument is, by nature of the logical argument, of the direct or indirect type. There will be an opportunity to discuss certain points of relation between the statistical inference theories according to Neyman-Pearson-Wald, often called the 'frequency theory', and according to the school of 'Bayesian theory'.

For the purpose of making inference with respect to the classificatory properties Q and $\bar{Q}$ defined in Section 6 we define the following random variables, namely, for the case of direct inference,

B_γ = The number of objects with property Q in the sample, given that the number of Q in the population is γ ,

and for the case of indirect inference,

Γ_β = The number of objects with property Q in the population, given that the number of Q in the sample is β .

The population and sample sizes are, as before, the fixed integers N and n , respectively. The following notation will be used,

$$P(B_\gamma = \beta) = f_\gamma(\beta), P(B_\gamma \leq \beta) = F_\gamma(\beta), P(B_\gamma \geq \beta) = F_\gamma^c(\beta),$$

$$P(\Gamma_\beta = \gamma) = g_\beta(\gamma), P(\Gamma_\beta \leq \gamma) = G_\beta(\gamma), P(\Gamma_\beta \geq \gamma) = G_\beta^c(\gamma),$$

where ' $P(\cdot)$ ' denotes the probability attached to the event within parenthesis.

The two probability distributions, the one for B_γ and the one for Γ_β , form a *distribution pair*, the first one being direct and the second one indirect in their interpretation through Carnap's theory. We shall observe how, given one of the distributions of the pair, the other one emerges as the result of a technical manoeuvre which may be called a transformation through Bayes' theorem. This involves the application of a suitable prior distribution. For example, denote the prior probability of the event that the population frequency of Q (denoted by ' Γ ') is equal to γ by

$$P(\Gamma = \gamma) = h(\gamma).$$

Then $f_\gamma(\beta)$ is transformed into $g_\beta(\gamma)$ by Bayes' theorem in the form

$$g_\beta(\gamma) = \frac{f_\gamma(\beta)h(\gamma)}{\sum_{\gamma} f_\gamma(\beta)h(\gamma)} \quad (7.1)$$

If statistical theory is dichotomised, as is often done, into 'frequency theory' and 'Bayesian theory', then the direct distribution $f_\gamma(\beta)$ is the one on which the frequency theorist would most likely rely for his inferences, while the adherent of Bayesian theory would be inclined to use the indirect distribution $g_\beta(\gamma)$.

In the direct inference case, take the degree of confirmation given by (6.3) and let it represent a probability distribution for B_γ , i.e.

$$f_\gamma(\beta) = \binom{\gamma}{\beta} \binom{\bar{\gamma}}{\bar{\beta}} / \binom{N}{n}, \quad (7.2)$$

where the range of values for β is given by (6.4).

In the indirect inference case, consider first the Carnapean formula (6.5), and let it represent a probability distribution for Γ_β , i.e.

$$g_\beta(\gamma) = \binom{\gamma + \omega - 1}{\beta + \omega - 1} \binom{\bar{\gamma} + \bar{\omega} - 1}{\bar{\beta} + \bar{\omega} - 1} / \binom{N + \kappa - 1}{n + \kappa - 1}, \quad (7.3)$$

for values of γ given by (6.6), and where the positive integers ω and $\bar{\omega}$ are the logical widths of Q and $\bar{Q}$, respectively, and $\omega + \bar{\omega} = \kappa$. We have

$$E\left(\frac{\Gamma_\beta + \omega}{N + \kappa}\right) = \frac{\beta + \omega}{n + \kappa} \quad (7.4)$$

where ' $E(\cdot)$ ' denotes the expected value of the random variable within parenthesis.

The distribution (7.3) is obtained from (7.2) through a transformation by Bayes' theorem (7.1) if the prior probability that the population frequency of Q equals γ is given by

$$P(\Gamma = \gamma) = h(\gamma) = \binom{\gamma + \omega - 1}{\omega - 1} \binom{\bar{\gamma} + \bar{\omega} - 1}{\bar{\omega} - 1} / \binom{N + \kappa - 1}{\kappa - 1} \quad (7.5)$$

for integers $\gamma, \bar{\gamma}$ such that $\gamma + \bar{\gamma} = N$. In particular, if $\omega = \bar{\omega} = 1$, then $h(\gamma)$ is a uniform distribution assigning equal probability to each combination $\gamma, \bar{\gamma}$ such that $\gamma + \bar{\gamma} = N$. We note that the prior distribution (7.5) is a particular case of $g_\beta(\gamma)$ as given by (7.3), namely, the one obtained when $n = \beta = 0$ in (7.3), i.e., when no sample data are available. Therefore, it follows by setting $n = \beta = 0$ in (7.4) that the prior expected proportion of objects possessing the property Q is

$$E\left(\frac{\Gamma}{N}\right) = \frac{\omega}{\kappa}.$$

Thus, speaking in terms of prior distributions, this is one interpretation of Carnap's (1952, p. 24) logical factor ω/κ .

Consider next the distribution obtained from (6.14) if we extend its area of application so as to include not only the cases $\delta = 0$ and $\delta = 1$, $\delta + \bar{\delta} = 1$. Instead, let us allow every pair of real numbers $\delta, \bar{\delta}$ such that

$$\delta \geq 0, \bar{\delta} \geq 0, \delta + \bar{\delta} = 1. \quad (7.6)$$

As binomial factors $\binom{a}{b}$ are usually defined only for integers a and b , we should write the resulting family of distributions in the form

$$g_\beta(\gamma) = \binom{N-n}{\gamma-\beta} \frac{I(\gamma+\delta, \bar{\gamma}+\bar{\delta})}{I(\beta+\delta, \bar{\beta}+\bar{\delta})}, \quad (7.7)$$

where the range of values of γ is given by (6.6), and where

$$I(a, b) = \mathcal{J}(a)\mathcal{J}(b)/\mathcal{J}(a+b),$$

and

$$\mathcal{J}(a) = \int_0^\infty x^{a-1} e^{-x} dx \quad (a > 0).$$

Provided $\beta + \delta > 0$, $\bar{\beta} + \bar{\delta} > 0$, formula (7.7) defines a probability distribution for Γ_β for any pair $\delta, \bar{\delta}$ satisfying (7.6). It has the form of a Pólya distribution, and we have

$$E\left(\frac{\Gamma_\beta + \delta}{N + 1}\right) = \frac{\beta + \delta}{n + 1}. \quad (7.8)$$

The distribution (7.7) is obtained from (7.2) through a transformation by Bayes' theorem (7.1) if we choose the prior probability that the population frequency of Q equals γ as

$$P(\Gamma = \gamma) = h(\gamma) = \binom{N}{\gamma} \frac{I(\gamma + \delta, \bar{\gamma} + \bar{\delta})}{I(\delta, \bar{\delta})} \quad (7.9)$$

for $\gamma = 0, 1, \dots, N$, and $\delta, \bar{\delta} > 0$, $\delta + \bar{\delta} = 1$. We note that (7.9) is a particular case of the transformed distribution $g_\beta(\gamma)$ given by (7.7), namely, the one obtained from (7.7) when no sample data are available, so that $n = \beta = 0$. To obtain the expected value of the distribution (7.9) we thus set $n = \beta = 0$ in (7.8). It follows that the prior expected proportion of objects possessing the property Q is

$$E\left(\frac{\Gamma}{N}\right) = \delta.$$

This gives an interpretation of the factor δ as the a priori estimated proportion of objects with property Q . The cases of $\delta = 0$, $\bar{\delta} = 1$ and $\delta = 1$, $\bar{\delta} = 0$ will be given special attention. (In the following discussion we always have $\bar{\delta} = 1 - \delta$.)

For $\delta = 0$, (7.7) takes the form (6.12), and $NQ = \nu$ in the argument of the preceding section. The adoption of (6.12) as a probability distribution for Γ_β can thus be said to represent the view that the population order of the object appearing as the last Q of the ordered sample sequence is equal to the population frequency of Q . For $\delta = 0$, (7.7) is obtained by a transformation of (7.2) by Bayes' theorem by giving prior probability proportional to γ^{-1} to the event that the population frequency of Q is γ .

For $\delta = 1$, on the other hand, (7.7) takes the form (6.13), and in the notation of the preceding section, $NQ = \mu - 1$. If (6.13) is taken as a probability distribution for Γ_β , it means that the object appearing as the first Q (from the left) of the ordered sample sequence is regarded as being of a population order that exceeds the population frequency of Q by as little as possible, i.e., by unity. For $\delta = 1$, (7.7) obtains through a trans-

formation of (7.2) by Bayes' theorem, assigning prior probability proportional to $\bar{\gamma}^{-1} = (N-\gamma)^{-1}$ to the event that the population frequency of Q is γ .

At best, both the of extreme cases $NQ = \nu$ and $NQ = \mu-1$ are decent approximations to the population number of Q . The true NQ usually lies somewhere between ν and $\mu-1$. This suggests formula (7.7), where δ lies between 0 and 1 (or possibly equals 0 or 1), as an indirect distribution representing the probability of the event $\Gamma_\beta = \gamma$.

In the theory of order statistics, formula (7.7) for $\delta = 0$ gives the direct probability distribution of a random variable representing the population order of the object of sample order β (see Wilks, 1962, p. 243). For $\delta = 1$, (7.7) gives the direct probability distribution of a random variable defined by subtracting unity from the population order of the object of sample order $\beta+1$.

Letting $\delta = 0$ in (7.7), the resulting distribution pair $f_\gamma(\beta)$ and $g_\beta(\gamma)$ fulfils certain simple relations, namely,

$$\beta f_\gamma(\beta) = \gamma g_\beta(\gamma), \quad (7.10)$$

$$F_\gamma^\circ(\beta) = G_\beta(\gamma), \quad (7.11)$$

$$E(B_\gamma) = \gamma E(B_1), E(\Gamma_\beta) = \beta E(\Gamma_1) \quad (7.12)$$

Formula (7.10) follows immediately from (7.2) and (7.7) with $\delta = 0$. Further, (7.11) is obtained by the following argument. The fact that $\delta = 0$ implies that if the population number of Q is γ , then γ is also the population order of the last Q of the ordered sample sequence, i.e. the object of sample order β . Make a classification of the population objects into those included in the sample and those not included in the sample. Then the first γ objects of the population ordered sequence will contain β or more sampled objects if and only if the population order of the β th sampled object (i.e. that of sample order β) does not exceed γ . Formula (7.12) may be shown to follow from (7.10) and (7.11) in combination.

The distribution pair obtained by (7.2) and (7.7) with $\delta = 1$ satisfies the same set of relations (7.10), (7.11), and (7.12) with $\bar{\gamma}$ and $\bar{\beta}$ substituted for γ and β , respectively, where $\bar{\gamma}$ and $\bar{\beta}$ are values taken by random variables $\bar{\Gamma}_\beta$, $\bar{B}_\gamma$ defined instead in relation to the property $\bar{Q}$.

8 A LIMIT CASE

Let N , γ , and $\bar{\gamma}$ approach infinity in such a way that $\gamma/N \rightarrow \theta$, $\bar{\gamma}/N \rightarrow \bar{\theta} = 1-\theta$ ($0 < \theta < 1$). The direct probability (7.2) then approaches the binomial distribution

$$f_\theta(\beta) = \binom{n}{\beta} \theta^\beta \bar{\theta}^{n-\beta}, \beta = 0, 1, \dots, n. \quad (8.1)$$

Under the same conditions, the indirect probabilities (7.3) and (7.7) approach limit distributions that have the form of beta densities, namely,

$$g_{\beta}(\theta) = \frac{1}{I(\beta + \omega, \bar{\beta} + \bar{\omega})} \theta^{\beta + \omega - 1} \bar{\theta}^{\bar{\beta} + \bar{\omega} - 1}, \quad 0 \leq \theta \leq 1, \quad (8.2)$$

and

$$g_{\beta}(\theta) = \frac{1}{I(\beta + \delta, \bar{\beta} + \bar{\delta})} \theta^{\beta + \delta - 1} \bar{\theta}^{\bar{\beta} + \bar{\delta} - 1}, \quad 0 \leq \theta \leq 1, \quad (8.3)$$

respectively. In (8.2), ω and $\bar{\omega}$ are positive integers such that $\omega + \bar{\omega} = \kappa$; in (8.3), δ and $\bar{\delta}$ satisfy (7.6). If $\delta = 0$ in (8.3), then the distribution pair formed by $f_{\theta}(\beta)$ and $g_{\beta}(\theta)$ will satisfy (7.10), (7.11), and (7.12), with θ substituted for γ .

The choice of a prior distribution gives Bayesian theory its air of arbitrariness. This is a parallel to the arbitrariness introduced into Carnap's theory by the choice of the primitive properties. The Bayesian theory connected with the binomial distribution offers an excellent example of the rich diversity of suggested prior distributions. We present a brief review of some suggested prior distributions. More detailed discussions of these matters are found in Good (1965, Section 3.2) and Plackett (1966).

A prior density belonging to the family of beta distributions

$$h(\theta) \propto \theta^{a-1} \bar{\theta}^{b-1}, \quad (8.4)$$

where a and b are real numbers, merges conveniently with the likelihood (8.1). This prior density yields, through Bayes' theorem, the posterior density

$$g_{\beta}(\theta) = \frac{1}{I(\beta + a, \bar{\beta} + b)} \theta^{\beta + a - 1} \bar{\theta}^{\bar{\beta} + b - 1}, \quad 0 \leq \theta \leq 1,$$

of which (8.2) and (8.3) are special cases.

The question of what values should be given to a and b in (8.4) is, of course, essential. Many suggestions, usually accompanied by some argument of justification for the choice, have been presented in the literature since the times of Bayes and Laplace. The suggestions for a and b are often produced by a desire of obtaining a result that is 'convenient' from a certain point of view. This kind of reasoning is usually quite removed from the type of logically coherent argument that underlies Carnap's theory, where the integers ω and $\bar{\omega}$ appearing in (8.2) could be identified as the logical widths of the underlying properties.

The suitability of the prior density (8.4) seems to have been first observed by Hardy (1889). The still older procedure of Bayes and Laplace consisted in assigning a uniform prior probability density for θ over $0 \leq \theta \leq 1$, expressing 'complete ignorance' about θ . This particular prior density is a special case of (8.4), namely, the one obtained if $a = b = 1$.

Fisher (1956) argued that the procedure of assigning a prior density is not consistent under transformations of the parameter. One might equally well assign a uniform prior density to any monotonic function of θ , whereby one would generally obtain an entirely different result. Fisher (1956, p. 16), in a discussion of the binomial case, considers the transformed parameter

$$\phi = \frac{1}{2} \cos^{-1} (1 - 2\theta).$$

A uniform distribution for ϕ corresponds to the prior density (8.4) with $a = b = \frac{1}{2}$.

Invariance theories have been constructed by Jeffreys (1961) and by Perks (1947), expressing a desire to present a rule for choosing the prior density under which consistent results are obtained under transformations of the parameter. In the case of the binomial distribution, Jeffreys' theory suggests the prior distribution (8.4) with $a = b = \frac{1}{2}$.

The particular prior density (8.4) with $a = b = 0$ has been advocated by Haldane (1945) and by Lindley (1965, vol. II, p. 145). Lindley notes that (8.4) is a proper prior density (i.e., integrable) if and only if $a > 0$, $b > 0$. The making of an observation will increase either a or b , and $a = b = 0$ gives the most desirable prior density since it represents the highest degree of ignorance to have a and b as small as possible. For $a = b = 0$, the resulting prior density is improper (i.e. the integral does not exist), but, technically speaking, this is not an obstacle in Bayesian theory.

Thatcher (1964) studied the properties of confidence limits in the frequency theory sense and in the Bayesian sense. In the case of the binomial parameter θ , he noted that the lower frequency theory limit equals the lower Bayesian limit if a prior probability density proportional to θ^{-1} is used, while the upper limits will coincide if a prior density proportional to $\bar{\theta}^{-1}$ is used. These prior densities are of the form (8.4) with $a = 1$, $b = 0$ and $a = 0$, $b = 1$, respectively.

We consider now the question of how to estimate the probability of 'success' (i.e. of obtaining an object with the classificatory property Q) in a future binomial trial.

The rule of maximum likelihood applied to (8.1) yields the estimate β/n . This is the mode of (8.2) for $\omega = \bar{\omega} = 1$, i.e. when the prior density is uniform. Good (1965) and Carnap (1950, p. 227) have discussed and criticised this estimate, which in Carnap's terminology is called the straight rule of estimation. Both authors note that difficulties arise with this rule if $\beta = 0$ or $\beta = n$. Interpreted in terms of betting odds, the case of $\beta = n$, for example, would result in estimating θ by unity, which implies willingness to bet an arbitrarily large sum in favour of a success at the next trial. This, of course, does not seem realistic.

Next, consider the estimates of θ provided by the expectation of θ in the posterior densities (8.2) and (8.3). These estimates are $(\beta + \omega)/(n + \kappa)$ and $(\beta + \delta)/(n + 1)$, respectively. Both of these estimates are free of the above mentioned serious drawback, since they do not attain the values 0 or 1 (provided $0 < \delta < 1$ in the latter estimate). In the estimate $(\beta + \omega)/(n + \kappa)$ we may, in Carnap's (1952, p. 24) terminology, distinguish β/n as an *empirical factor* and ω/κ as a *logical factor*.

The estimate $(\beta + \delta)/(n + 1)$ is also determined by two factors, namely by the empirical factor β/n and by δ . While the estimate $(\beta + \omega)/(n + \kappa)$ may deviate substantially from the straight rule estimate β/n , the absolute value of the difference between β/n and $(\beta + \delta)/(n + 1)$ will always be small; it is easily shown to be less than $(n + 1)^{-1}$ for all β and δ , $0 < \delta < 1$. The straight rule estimate β/n is therefore never 'distorted' to any great extent by the presence of the factor δ in the estimate $(\beta + \delta)/(n + 1)$. If the straight rule estimate is rejected solely on the grounds that it can take on the values 0 and 1, the procedure $(\beta + \delta)/(n + 1)$ provides an estimate free from this weakness (if $0 < \delta < 1$), but it is still approximately equal in value to β/n . Arbitrariness, represented in Carnap's approach by the factor ω/κ , is present here in guise of the factor δ . A compromise is to choose the midpoint value of $\delta = 1/2$ (cf. Good, 1965, p. 18).

Further discussions with a bearing on this topic are found in Carnap's (1952) work on the 'continuum of inductive methods'.

9 NOTES ON STATISTICAL PROCEDURE

In conclusion, we shall point out a few examples where the results of this paper have a bearing on statistical procedure.

Thatcher (1964) and other authors referred to in that paper have investigated the conditions under which confidence intervals (in the frequency theory sense) for a certain parameter agree with the Bayesian intervals in the same situation. The establishment of such agreements is a result of identity relations of the type (7.11).

Here, we consider as an example the case where the random variables B_γ and Γ_β are distributed according to (7.2) and (7.7), respectively, where we take $\delta = 0$ in (7.7). As we have observed previously, this is equivalent to attaching prior probability proportional to γ^{-1} to the event that the population frequency of Q is γ . Under these conditions, (7.11) holds.

Let the integer γ_e be such that

$$G_\beta(\gamma_e) = \alpha > 0.$$

Then

$$P(\gamma_e < \Gamma_\beta) = 1 - G_\beta(\gamma_e) = 1 - \alpha$$

so that $\gamma_e < \gamma$ is a one-sided Bayesian interval for γ at the level $1 - \alpha$.

However, because of (7.11) we also have

$$F_{\gamma_0}^{\circ}(\beta) = \alpha.$$

Let γ_0 be a fixed value of γ . Then $\gamma_c < \gamma_0$ if and only if β is such that

$$\alpha = F_{\gamma_0}^{\circ}(\beta) < F_{\gamma_c}^{\circ}(\beta), \quad (9.1)$$

for $F_{\gamma}^{\circ}(\beta)$ is (stepwise) increasing as γ increases, for a fixed β . The inequality (9.1) holds for all β such that $\beta \leq \beta_0$, where β_0 is the largest β for which $F_{\gamma_0}^{\circ}(\beta)$ exceeds α , because $F_{\gamma}^{\circ}(\beta)$ decreases as β increases, for a fixed γ . Thus $P(B_{\gamma_0} > \beta_0) \leq \alpha$, and $\gamma_c < \gamma_0$ if and only if $B_{\gamma_0} \leq \beta_0$. Therefore, with the frequency theory interpretation of confidence intervals, by which γ_c is a random variable,

$$P(\gamma_c < \gamma_0) = P(B_{\gamma_0} \leq \beta_0) = 1 - P(B_{\gamma_0} > \beta_0) \geq 1 - \alpha.$$

Thus, the lower one-sided interval $\gamma_c < \gamma$, shown above to be a Bayesian interval at level $1 - \alpha$, is at the same time a frequency theory confidence interval at a level no less than $1 - \alpha$.

For the upper one-sided interval $\gamma > \gamma_c$, a similar result is obtained if we take $\delta = 1$ in (7.7), which corresponds to giving prior probability proportional to $\bar{\gamma}^{-1} = (N - \gamma)^{-1}$ to the event that the population frequency of Q is γ .

Our final remark concerns the theory of statistics relating to inference from observed observations on the one hand and grouped samples on the other. Consider a grouped sample of size n such that β observations are known to fall in the interval $-\infty < x \leq x_0$ and $\bar{\beta} = n - \beta$ observations in the interval $x_0 < x \leq \infty$. Let $\theta = F(x_0)$, $\bar{\theta} = 1 - F(x_0)$, where $F(x)$ is assumed to be the known, continuous distribution function of the random variable X on which the grouped sample constitutes n independent observations. (The assumption of a continuous distribution is equivalent to assuming an infinite population.) The distribution of B_{γ} , the number of individuals in the lower interval, is binomial and is by a direct argument given by (8.1).

Consider now for the same situation an indirect argument in the vein of the distribution (8.3). The random variable in this connection can be taken as $\theta = F(x_0)$, or as x_0 itself, since x_0 is related to θ through a transformation $F(x)$ of known form. The class limit x_0 is such that $X_{(\beta)} \leq x_0 < X_{(\beta+1)}$, where $X_{(i)}$ denotes the i th largest observation of the ordered sample. Formula (8.3) can also be interpreted as a direct probability distribution in the theory of order statistics from an infinite population with the distribution function $F(x)$. For $\delta = 0$ it represents the density function for $\theta = F(X_{(\beta)})$, and for $\delta = 1$, it represents the density function for $\theta = F(X_{(\beta+1)})$. Now, every class limit x_0 located in the interval $X_{(\beta)} \leq x_0 < X_{(\beta+1)}$ would have produced the same class frequencies β and $\bar{\beta}$.

In view of these considerations, an argument can be raised in favour of letting the indirect distribution of x_0 be such that the density of $\theta = F(x_0)$ is given by (8.3) for some δ , $0 < \delta < 1$. Again a compromise choice is $\delta = \bar{\delta} = \frac{1}{2}$.

University of Umeå

REFERENCES

- Carnap, R. (1950), *Logical Foundations of Probability* (Chicago: Chicago University Press).
- Carnap, R. (1952), *The Continuum of Inductive Methods*. (Chicago: Chicago University Press).
- Carnap, R. and Stegmüller, W. (1959), *Induktive Logik und Wahrscheinlichkeit* (Vienna: Springer-Verlag).
- Good, I. J. (1965), *The Estimation of Probabilities. An Essay on Modern Bayesian Methods* (Cambridge, Mass.: MIT Press).
- Hardy, G. F. (1889), correspondence in *Insurance Record* (1889), reprinted in *Trans. Fac. Actuaries*, 8 (1920).
- Fisher, R. A. (1956), *Statistical Methods and Scientific Inference*. (London: Oliver and Boyd).
- Jeffreys, H. (1961), *Theory of Probability*. 3rd edn. (Oxford: Clarendon Press).
- Lindley D. V. (1965), *Introduction to Probability and Statistics, Part I: Probability Part II: Inference* (Cambridge: Cambridge University Press).
- Perks, W. (1947), 'Some Observations on Inverse Probability Including a New Indifference Rule', *J. Inst. Actuar.* 73, 285-312.
- Plackett, R. L. (1966), 'Current Trends in Statistical Inference', *J. R. Statist. Soc. A*, 129 249-67.
- Thatcher, A. R. (1964), 'Relationships between Bayesian and confidence limits for prediction', *J. R. Statist. Soc. B*, 26, 176-92.
- Tintner, G. (1949), 'Foundations of Probability and Statistical Inference', (with discussion). *J. R. Statist. Soc. A*, 112, 251-86.
- Wilks, S. S. (1962), *Mathematical Statistics* (New York: Wiley).

Discussions

THE JONES CASE

In a recent article [1] Isaac Levi has proposed two statistical counter-examples to Richard Jeffrey's method of probability kinematics. Although Levi makes other points against Jeffrey, some of which are well taken, neither of these counter-examples is successful. We shall first give as much background about probability kinematics and Jeffrey's system as is necessary, then we shall present and refute Levi's counter-examples.

I

Probability kinematics may be regarded as the study of how an agent ought to revise his probability function in response to new experience. If at a time t_0 a person S attributed various probabilities to the propositions in a set K , then S 's probability function at t_0 would be the function which assigned to each proposition H in K the probability that S attributed to H . If at a later time t_1 S gains new information relevant to the probabilities he attributes to some member of K , then S ought to make appropriate revisions in his probability function for the rest of the members of K . Jeffrey has developed a method for determining what revisions ought to be made in one's probability function throughout K , given that he has modified his probability assignments in some subset of K ([2] ch. 11). Jeffrey's method does not specify how one ought to make the initial modifications; all it does specify is how the rest of the probability assignments in K ought to be revised, once such a shift has taken place.

Conditionalisation is a standard method for specifying such revisions when the initial modification is that of accepting some proposition E in K as evidence, i.e. of assigning to E a new probability of 1. If $Prob$ is S 's initial probability function for K , and if S modifies his probability assignment to E from $Prob(E)$ to a new probability assignment $Pr_{ib}(E)$ where $Pr_{ib}(E) = 1$; then, for all H in K , S ought to assign to H a revised probability $Pr_{ib}(H)$ such that

$$Pr_{ib}(H) = \frac{Prob(HE)}{Prob(E)}$$
 If Pr_{ib} is S 's probability function when correctly revised in

response to accepting E as evidence, then, for all H in K , $Pr_{ib}(H)$ is the conditional probability of H on E , which will be written $Prob(H/E)$.

Jeffrey finds conditionalisation inadequate as a general method for adjusting $Prob$ to an appropriate Pr_{ib} in response to some shift in probability assignments in a subset of K because it will only work when the original shift is that of assigning, to some E in K , $Pr_{ib}(E) = 1$ ([2] p. 154). His own method covers all cases where the original shift is from $Prob(E)$ to $Pr_{ib}(E) < 1$, and it is equivalent to conditionalisation for the cases where conditionalisation works ([2] p. 154).

It is reasonable to suppose that there may be actual cases worth bothering about where conditionalisation would be inadequate when one considers observations made under poor perceptual conditions. On the basis of this sort of observation one might be justified in changing his probability assignment to a proposition even though he would not be justified in accepting either the proposition or its negation as evidence. Jeffrey provides an example where a subject is trying to tell the colour of a piece of cloth in poor light ([2] p. 154). A more interesting sort of case where conditionalisation might fail sometimes occurs during space shots. The control group may receive information on which they must change their probability assignment to the proposition that some part will malfunction from 0.001 to 0.04. In such a case it is of the utmost importance for them to reassess the probability assignment to the proposition that the shot will be successful. Even a small increase in the probability of failure of some crucial component could be good grounds for terminating the shot, especially if it is manned, while a rather large increase in the probability of failure of some relatively peripheral component might be tolerated. Jeffrey's probability kinematics provides a useful and straightforward device for calculating the changes in the failure-probability of one component of a system, given a change in failure probability of another component. It is worth noting that there need be nothing subjectivistic about these probabilities in order for the theory to apply.

Jeffrey's method is the following. Consider any case where S changed his probability assignment to E in K from $Prob(E)$ to $Prib(E)$. It is a principle of Jeffrey's system that no such change will alter the conditional probability of H on E for any H in K . Thus, whenever the original shift is in the probability assignment to E , then

$$(1) \quad Prob(H/E) = Prib(H/E) \quad ([2] \text{ p. 158})$$

This principle, together with the general coherence requirements on any probability function, is sufficient to deduce the following rule: For all H and K if $Prob(E)$ has changed to $Prib(E)$, then

$$(2) \quad Prib(H) = Prob(H/E) Prib(E) + Prob(H/\bar{E}) Prib(\bar{E}) \quad ([2] \text{ p. 159})$$

This rule can be generalised to cover any coherent case where the originating change is in any subset of up to any number n of members of K ([2] pp. 161-4). For our discussion, however, Rule (2), as given, will be sufficient.

2

Levi's first counterexample is a case where at a time t_0 a subject Jones is confronted with an urn which he knows contains either 5 blue and 5 green discs, or 10 blue and 0 green discs. Jones also knows that one disc is to be drawn from the urn. Let us call this background information, available to Jones at t_0 , b . Let H be the proposition that the urn contains 5 blue and 5 green discs; then, given b , not- H is equivalent to the proposition that the urn contains 10 blue and 0 green discs. Let E be the proposition that the disc drawn is blue; then, given b , not- E is equivalent to the proposition that the disc drawn is green. Thus (symbolising not- H by ' $\bar{H}$ ' and not- E by ' $\bar{E}$ ') on his background information Jones is justified in claiming that $H \vee \bar{H}$ and $E \vee \bar{E}$ are both logically true, and should each be assigned probability 1, and that both $E\bar{E}$ and $H\bar{H}$ are contradictory and should be assigned

probability 0. At t_1 Jones has an opportunity to observe the disc drawn and revises his probability assignment to E from 0.6 to 0.9. Let *Prob* be Jones's probability function at t_0 and *Prib* be the probability function he ought to have in response to the change in his probability assignment to E ; then *Prob* (E) = 0.6 and *Prib* (E) = 0.9.

According to Levi it should be the case that *Prob* (E/H) = *Prib* (E/H) = 0.5 for Jones. There is no problem in granting that *Prob* (E/H) ought to equal 0.5, for given only the background information at t_0 with the sole modification that H is accepted, one would be justified in claiming that the disc would be drawn from an urn containing 5 blue and 5 green discs and would have no other information relevant to E . Therefore, it would be most reasonable for Jones to assign *Prob* (E/H) = 0.5. Levi justifies his claim that *Prib* (E/H) should also equal 0.5 as follows ([1], p. 206):

Having admitted no new evidence at time t_1 , Jones at that time ought surely to consider the draw from the urn to be a 'random' one. *Prib* (E/H) ought surely to equal *Prob* (E/H).

Having established to his satisfaction that *Prob* (E/H) = *Prib* (E/H), Levi points out that on Jeffrey's principle *Prob* (E/H) cannot consistently equal *Prib* (E/H) in the Jones case, or in any case where *Prob* (E) $\neq$ *Prib* (E) ([1], p. 206). Therefore, he finds Jeffrey's system in conflict with statistical good sense.

The authors have succeeded in deriving the contradiction claimed by Levi (see appendix), but Levi's statement seems to suggest that he had something simpler in mind than the derivation we present. One possibility is the following: since no new evidence has been acquired, *Prob* (H) ought to be equal to *Prib* (H); or if it be maintained that the glimpse of the disc under poor light is after all *some* evidence, nevertheless *Prib* (H) and *Prob* (H) ought surely to be very close in value. But, applying the general multiplication theorem, this is to say that

$$(3) \quad \frac{\textit{Prib} (E) \cdot \textit{Prib} (H/E)}{\textit{Prib} (E/H)}$$

ought to be identical to, or very close in value to

$$(4) \quad \frac{\textit{Prob} (E) \cdot \textit{Prob} (H/E)}{\textit{Prob} (E/H)}$$

Now Jeffrey's rule requires us to take *Prib* (H/E) = *Prob* (H/E), and Levi argues that we should take *Prib* (E/H) = *Prob* (E/H); but this, combined with the assumption that *Prib* (H) and *Prob* (H) are identical or close in value, conflicts with our supposition that *Prib* (E) and *Prob* (E) were significantly different.

What are we to say to this counterexample? The case turns out exactly as it should, for *Prob* (E/H) ought not to equal *Prib* (E/H)! At time t_1 the background information has been changed in a way that directly effects the probability assignment to E . Though there has been no new evidence accepted, nevertheless, at time t_1 Jones has observed the disc sufficiently to be justified in assigning *Prib* (E) = 0.9. Therefore, given this background information and the assumption that H is accepted, it would not be reasonable to assume the probability of E = 0.5. But this is exactly what one must assume when he holds *Prib* (E/H) = 0.5, and this precisely and plausibly contradicts Levi's claim that the draw is a

'random' one. Indeed, we can specify why it is not a random member of the set of discs drawn from the urn with respect to being blue: it is because we know *more* about the disc than that it has been drawn from the urn; we know that, glimpsed fleetingly, it appears blue. This glimpse suffices to justify our assigning a probability of 0.9 to the proposition that it is blue. And the glimpse thereby allows us to go beyond the bare statistical information provided by H and $\bar{H}$.

It is important to notice that Jeffrey's principle (1) does not run into this difficulty. Given the new background information which includes the observation on which one assigns $\text{Pr}ib(E) = 0.9$, and also adding the assumption that E is evidence, one ought to assign to H the same probability one would assign to H given only the old background information together with the assumption that E is evidence. This follows because the assumption that E is accepted, which is part of $\text{Pr}ib(H/E)$, renders the observation redundant.

Levi's second counter example is obtained from the Jones case merely by changing $\bar{H}$ to the hypothesis that there are 7 blue discs and 3 green discs in the urn ([1], p. 206). This case is actually another way of showing that on Jeffrey's method Levi's assumption that $\text{Prob}(E/H) = \text{Pr}ib(E/H)$ cannot hold for Jones. It is worth examining it separately because the example reveals an interesting virtue of Jeffrey's Rule 2. Rule 2 follows from Jeffrey's principle that $\text{Prob}(H/E) = \text{Pr}ib(H/E)$; in the same way, Rule 2' follows from Levi's principle that $\text{Prob}(E/H) = \text{Pr}ib(E/H)$.

$$(\text{Rule } 2') \text{Pr}ib(E) = \text{Prob}(E/H) \cdot \text{Pr}ib(H) + \text{Prob}(E/\bar{H}) \cdot \text{Pr}ib(\bar{H})$$

Substituting the values 0.9 for $\text{Pr}ib(E)$, 0.5 for $\text{Prob}(E/H)$ and 0.7 for $\text{Prob}(E/\bar{H})$, as the new Jones case requires, we have:

$$\begin{aligned} 0.9 &= 0.5 \text{Pr}ib(H) + 0.7 \text{Pr}ib(\bar{H}) \\ 0.9 &= 0.5 \text{Pr}ib(H) + 0.7 (1 - \text{Pr}ib(H)) \\ 0.9 &= 0.7 - 0.2 \text{Pr}ib(H): \text{impossible.} \end{aligned}$$

A contradiction! This result is exactly as it should be, for R2' specifies the change Jones ought to make in his probability assignment to E in response to a shift in his probability assignment to H . If Jones had changed his probability assignment for H to 0, thereby accepting $\bar{H}$ as evidence, then he should assign 0.7 as the new probability of E . 0.7 is thus the highest value for $\text{Pr}ib(E)$ that could be justified by a change in Jones's probability assignment to H . Therefore, as the contradiction indicates, the high $\text{Pr}ib(E)$ value of 0.9 could not have resulted from Jones's rational response to any shift in his probability assignment to H . $\text{Pr}ib(E)$ must originate in some other source; and that source is, of course, the observation.

APPENDIX

A general proof that it cannot be the case both that $\text{Prob}(E/H) = \text{Pr}ib(E/H)$ and that $\text{Prob}(E) \neq \text{Pr}ib(E)$ is the following:

- | | | |
|---|--|---------------------|
| 1 | Premise: $\text{Pr}ib(H/E) = \text{Prob}(H/E)$ | Jeffrey's Principle |
| 2 | Premise: $\text{Pr}ib(E/H) = \text{Prob}(E/H)$ | Levi's Principle |
| 3 | Premise: $\text{Pr}ib(E/\bar{H}) = \text{Prob}(E/\bar{H})$ | Levi's Principle |

Multiplication theorem of Probability (coherence):

$$4 \quad \text{Pr}ib(H) = \frac{\text{Pr}ib(E) \cdot \text{Pr}ib(H/E)}{\text{Pr}ib(E/H)}$$

$$5 \quad \text{Pr}ib(\bar{H}) = \frac{\text{Pr}ib(E) \cdot \text{Pr}ib(\bar{H}/E)}{\text{Pr}ib(E/\bar{H})}$$

Addition theorem, principle of total probability:

$$6 \quad \text{Pr}ib(H \vee \bar{H}) = \text{Pr}ib(H) + \text{Pr}ib(\bar{H}) = 1$$

(4, 5, 6):

$$7 \quad 1 = \text{Pr}ib(E) \left(\frac{\text{Pr}ib(H/E)}{\text{Pr}ib(E/H)} + \frac{\text{Pr}ib(\bar{H}/E)}{\text{Pr}ib(E/\bar{H})} \right)$$

Analogously:

$$8 \quad 1 = \text{Prob}(E) \left(\frac{\text{Prob}(H/E)}{\text{Prob}(E/H)} + \frac{\text{Prob}(\bar{H}/E)}{\text{Prob}(E/\bar{H})} \right)$$

(1, 2, 3, 7):

9 substitute *Prob* for *Pr*ib in right hand side of 7

(7, 9):

$$10 \quad \text{Prob}(E) = \text{Pr}ib(E)$$

WILLIAM L. HARPER
HENRY E. KYBURG
University of Rochester

REFERENCES

- [1] Isaac Levi, 'Probability Kinematics', *Brit. J. Phil. Sci.* 18 (1967), 197-209.
[2] Richard C. Jeffrey, *The Logic of Decision* (New York, 1965).

A NOTE ON TWO 'PROBLEMS' OF INDUCTION

(1) I may be quite naive—but it seems to me that some more recent 'problems' of induction are no problems at all, but can be solved in a line or two.

(2) To see this, let us first introduce the general notion of an *instance* of a hypothesis. Given a hypothesis

$$(x)[Ax \rightarrow Bx] \quad (i)$$

we may define various types of instances, viz.

- positive instances: all those individuals which are both *A* and *B* ;
- negative instances: all those individuals which are both *A* and $\bar{B}$;
- disjunctive instances: all those individuals which are either *A* or *B* ;
- and so on.

Considering that there are sixteen truth functions of two arguments we arrive at sixteen (or, omitting tautologies and contradictions, fourteen) different kinds

of instances. Shall we use *all* of them in our discussion of scientific method and, if not, which instances shall we omit?

It is clear that we shall use only those instances which enable us to make judgments either about the *truth value*, or at least about the *probability* of the corresponding hypothesis. They are the only important ones, and only their behaviour is of interest.

Thus the fact that the disjunctive instances show the 'paradoxical' property of giving different results for different formulations of the same hypothesis, viz.

$$\begin{array}{ll} (x)[Ax \rightarrow Bx] & \text{disj. Instance: } A_j \vee B_j \\ (x)[\bar{B}x \rightarrow \bar{A}x] & \text{disj. instance: } \bar{A}_j \vee \bar{B}_j \end{array}$$

will not worry us a bit *unless* it emerges that they are relevant in the sense just explained and may therefore say different, and perhaps even incompatible things about the same hypothesis. Are positive instances relevant instances?

(3) Hume's arguments show that they are not: given individuals which are both A and B we cannot infer anything concerning the truth value, or even concerning the probability (relative frequency) of (i).¹ It is of course always possible to introduce functions of positive instances whose value increases with the number of positive instances found. It may even be the case that some men or beasts behave in accordance with functions of this kind, i.e. believe strongly when the function shows a high value, and start becoming doubtful, and downright nervous when the function gives a low value. But to make such a function the basis of scientific method would be quite unsatisfactory. First, because not everyone reacts in the manner indicated. There are people who distrust success and who start smelling a rat when a hypothesis moves from confirmation to confirmation with never a difficulty in sight (it would also seem that such people do not at all react irrationally but can give very good reasons for their behaviour). Secondly, constructing functions of the kind described would mean evaluating hypotheses by common consent, or by arbitrary stipulation favouring what is old and familiar. Thirdly, the preservation of 'popular' functions in the past would most certainly have prevented the overthrow of the Ptolemaic-Aristotelian cosmology, of the belief in the existence and omnipresence of witchcraft, and of similar 'backward' ideas.² Fourthly—and this applies especially to the positive instances—because there are cases where the functions give contradictory results.

Considering such objections it would seem to be desirable to eliminate all such functions and to use the laws of logic only. This eliminates all but the *negative instances*.

¹ That transition to (objective) probabilities does not improve matters is argued by Hume himself in the anonymous pamphlet *An Abstract of a Book lately published entitled a Treatise of Human Nature*, 1740, which was rediscovered by Keynes and Straffa and published by them in 1938. Cf. Appendix A of McNabb's edition, Meridian Books, New York 1962.

² Cf. my 'Problems of Empiricism' in *Beyond the Edge of Certainty*, ed. R. S. Colodny (Prentice Hall, 1965) as well as 'Problems of Empiricism, Part ii' to be published in one of the future volumes of the *Pittsburgh Studies in the Philosophy of Science*. The latter essay contains historical material on the fate of the Copernican hypothesis at the time of Galileo.

(4) Now it can be shown¹ that negative instances suffice in science. Of course, they do not explain why a certain hypothesis succeeds, for example, they do not explain why, given a place not too far north, the sun always rises (approximately) in the east. But this is not a task for instances anyway. A task of this kind is carried out by (more general) hypotheses such as the law of gravitation. Negative instances, therefore, are not only the sole survivors of a closer look at the relation between hypotheses and the 'evidence', they are also entirely sufficient for methodology. Let us now see what their acceptance (to the exclusion of all other instances) does to the new 'problems' of induction.

(5) The *paradox of confirmation* consists in the fact that the positive instances of a hypothesis depend on its formulation: 'all ravens are black' which may also be formulated 'all non-black things are not ravens' is supported, in the first case, by black ravens, in the second case by green peas, red noses etc. As can be seen from the list, the negative instances do not exhibit this variation (a property that can be proved to hold quite generally). But as they are the only relevant instances, there is no paradox to be worried about.

FORMULATION	POSITIVE INSTANCES	NEGATIVE INSTANCES
$(x)[Ax \rightarrow Bx]$	$A_i \ \& \ B_i$	$A_i \ \& \ \bar{B}_i$
$(x)[Bx \rightarrow Ax]$	$\bar{A}_i \ \& \ \bar{B}_i$	$A_i \ \& \ \bar{B}_i$
$(x)[Ax \vee \bar{A}x \rightarrow Ax \vee Bx]$	$A_i; \ B_i$	$A_i \ \& \ \bar{B}_i$

Goodman's paradox may be put as follows: compare the hypothesis

$$(x)[Rx \rightarrow Bx] \text{ (all ravens are black)} \quad (\text{ii})$$

with the hypothesis

$$(x)[Rx \rightarrow B'x] \text{ (all ravens are blite)} \quad (\text{iii})$$

where the predicate 'blite' is defined as 'black, if found before midnight tonight, white otherwise'. Assume that all ravens found until now are black (and therefore also blite). Now according to the common way of looking at the matter (common among experts, not among normal people!) hypothesis (iii) will be strongly confirmed, but at the same time we are invited to disregard it and to choose (ii) instead (black ravens found make it very unlikely that a few hours from now we shall find white ravens). This is as clear a refutation as can be desired of the idea that a hypothesis can be confirmed, and that the confirmatory force lies in its positive instances. There is again no trouble if we consider negative instances only.

(6) In view of the above results, of Hume's arguments, of the fruitfulness of the point of view of falsification, is it really asking too much to terminate the discussion of such red herrings as these and to concentrate upon the more rewarding task of advancing knowledge by the criticism of old hypotheses and the suggestion of new ones?

PAUL K. FEYERABEND
London-Berkeley-Berlin

¹ Cf. K. R. Popper, *Logic of Scientific Discovery*—if we omit the somewhat strange theory of corroboration.

SIMULTANEITY AND THE FUTURE

In his paper 'Time and Physical Geometry',¹ Professor Hilary Putnam argues for the 'reality' of future events as follows: He enunciates three premisses the last of which he identifies with the principle that there are to be no privileged observers. These are:

1. I-now am real. (I shall call myself O_1)
2. At least one other observer is real, and it is possible for this other observer to be in motion relative to me. (I shall call this observer O_2)
3. If it is the case that all and only the things (events) that stand in a certain relation R to me-now are real, and you-now (the other observer) are also real, then it is also the case that all and only the things that stand in the relation R to you-now are real.

He then presumes that you (O_2) are passing me (O_1) let us say at point P-now at a velocity significant compared to that of light, and claims that some events which are future relative to O_1 are present (stand in the relation R) to O_2 , and accordingly that at least some future events are real.

Let this be illustrated on his Fig. 1.

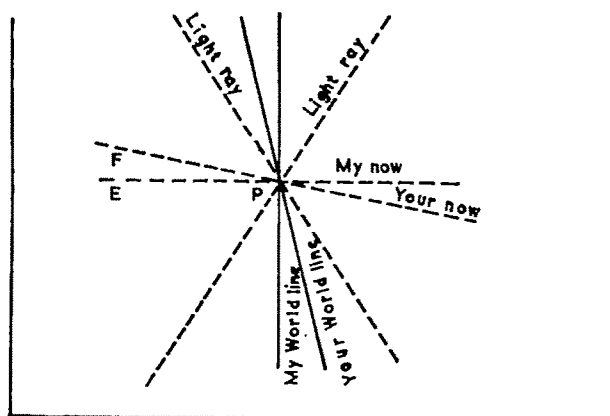


Fig. 1.

Event F is marked on the line which Professor Putnam calls 'your now'. Event E on the line called 'my now', and F is supposed to be future to me but present to you.

There is in this argument at least one assumption inconsistent with the Special Theory of Relativity on which the reasoning claims to base itself. It is assumed (a) that all events on the line my-now are in the relation R to O_1 and those on the line you-now are in the relation R to O_2 and (b) that R holds on the line 'me-now' in virtue of the simultaneity with P of all events on that line for O_1 , and similarly on the line 'you-now' for O_2 . But whether or not (a) is true, (b) is indeterminate at P. For according to special relativity the only events

¹ *The Journal of Philosophy* 64 (1967), 8.

simultaneous with O_1 and O_2 at P are the events occurring at P (i.e. in the here-now). All events outside the light cone (represented by the dotted lines marked 'light ray') are absolutely elsewhere, and, because simultaneity at a distance is not operationally determinable by O_2 , or by O_1 at P,¹ it is impossible to say whether they are present, past or future in relation to P. The 'now' lines in the diagram are therefore illegitimate and misleading.

It does not help to define the relation R as 'the relation of simultaneity in the observer's co-ordinate system' (as Putnam does) because no events are determinably simultaneous with P-now except those occurring at P-now, and which distant events any observer might calculate to have been simultaneous in his co-ordinate system he can determine only at some later time. Accordingly, what distant events bear the relation R to me-now is unknowable to me-now and becomes known to me only later. At some later date I shall be able to judge that E was simultaneous with P in my co-ordinate system and that F occurred later, and you will be able to calculate that F was simultaneous with P in your co-ordinate system; but by that time both E and F will be past in relation to both observers. So the truth value of the propositions 'E is now real' and 'F is now real' would both be indeterminable at P. It would seem odd to argue that some events which are future to me-now at P are now-real because when they become past to both you and me we shall discover that they came after events which I shall then take to have been simultaneous with P. It would be more natural to say that at any one moment no observer can determine what events (other than those spatially located at P) are simultaneous with him at that moment, or whether they are now-real.

Accordingly, it is not possible at P to say that F is in the future for O_1 or in the present for O_2 ; these statements are indeterminate. For observers at P no statements about such events are verifiable because no signals from them can reach observers at P. Whether on that account one would wish to say that such statements have no truth value is an open question exactly parallel to Aristotle's about the sea-battle.

Professor Putnam admits that the relation R cannot be 'simultaneity-in-my-co-ordinate-system', for then it would not be transitive as is required by the principle that there are no privileged observers. But he goes on to assume (at least so far as I can understand him) that if R satisfies the condition that all things that exist *now* are real, and is transitive, then everything that exists now in my co-ordinate system as well as in yours is real, and that some of these things will be future to one or other of us. But that does not follow for the reason given above—that in any co-ordinate system only what exists here-and-now is really

¹ If t_0 is the time of me-now at P, and the event E (at a distance) is the emission of a light signal, I cannot say: 'E is simultaneous with me-now if I shall receive the signal at $t_0 + \frac{x}{c}$ ' (where x is the distance of E from P and c the velocity of light), because x can be measured only if it is assumed that E and P are simultaneous. The definition would thus be circular. (Cf. Eddington, *The Philosophy of Physical Science*, Cambridge, 1939, p. 38; and Emile Borel, *Space and Time*, Dover Publications, New York, 1960, p. 155.)

No distant event can be observed at the time of its occurrence. When observed all events are already past, and among them those judged to have been simultaneous will differ in different co-ordinate systems.

present and what exists absolutely elsewhere cannot be specified at P. Special relativity, so far from decreasing the indeterminacy of the future extends it likewise to the distant.

Professor Putnam objects that if Oscar is a person whose whole world line is outside the light cone relative to P then at a future stage of O_1 , when Oscar is in the lower half of O_1 's light cone, it will be true that Oscar existed, but it will never be true that Oscar exists now. This is not strictly correct. 'Oscar exists' may be true at some time but it would not be possible for O_1 to say when it was true until after Oscar had ceased to exist. But even this is not strictly correct, for only a privileged observer (presumably God) could say whether Oscar had ceased to exist at P_1 (see Fig. 2), though at some later date in his time scheme O_1 could calculate whether or not it was true.

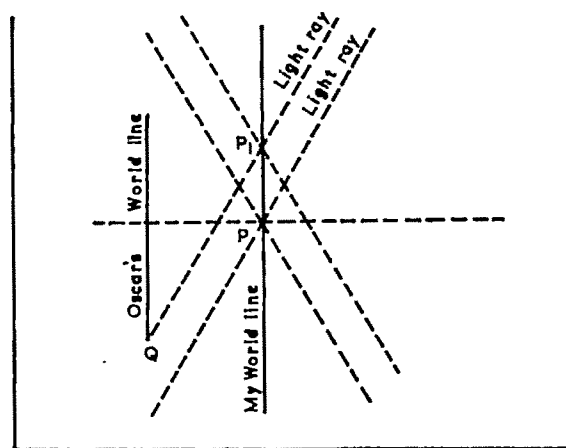


Fig. 2.

The problem of the determinacy of the future is thus unresolved so long as the principle is maintained that no observer is privileged. At P Oscar's existence is unknowable to me (O_1). At P_1 I know that he existed but not whether he still exists. It is, of course known to Oscar at Q that he exists, but it does not make sense to ask whether Q is present, past or future to P. When they discover it O_2 will date Oscar's existence differently from O_1 , so the statement 'Oscar exists at time t ' will have different truth values for O_1 and O_2 and for all other observers (including Oscar) moving relative to one another. Events cannot be *simply* located in space or time, for each observer has his own local time.

The 'reality' or determinability of events depends not only on their futurity, presentness or pastness, but also on their relative position in space.

Further, even if Professor Putnam's argument were admissible, it would not affect the 'reality' or determinability of events in the upper portion of the light cone—the absolute future—of any observer at P.

ERROL E. HARRIS
Northwestern University

A NOTE ON GOODMAN'S PARADOX

It has sometimes been thought that Goodman's paradox might be solved if a solid distinction, not relative to a particular language, could be made between qualitative and non-qualitative predicates, according to which Goodman's predicate 'grue' would be non-qualitative, while 'green' would be qualitative. For once such a distinction were available, one might reasonably lay it down that qualitative predicates should be projected in preference to non-qualitative predicates. The point of this note is to show that while this may be one point to make, it cannot be *the* fundamental point of the paradox, as R. G. Swinburne [1] has recently claimed. There is a purely qualitative predicate which is paradoxical in just the way that 'grue' is paradoxical. The predicate in question is incidentally also an observational predicate, so that the distinction between observational and non-observational predicates has no decisive bearing on the paradox.

Suppose that 99 is the number one greater than the largest number of facets on any examined emerald. Then consider the predicate 'gracet', defined as follows:

(A) x is gracet if and only if *either* x has fewer than 99 facets and is green *or* x has at least 99 facets and is blue.

It should be immediately clear that on any reasonable criterion 'gracet' is purely qualitative. For instance, on a criterion proposed by Swinburne [1], 'gracet' is qualitative. For the primary tests of whether something is gracet do not require finding out its spatial or temporal relation to any particular object, event, place or time. 'Gracet' is also relevantly similar to Goodman's 'grue'. For

- (1) Many examined emeralds are gracet.
- (2) No non-gracet emerald has been observed.
- (3) There are some as yet unexamined emeralds that may be gracet.
- (4) Projection of the hypothesis 'All emeralds are gracet' is (intuitively) unreasonable.
- (5) Just as 'green' can be 'defined' in terms of 'grue' and 'bleen', so can 'green' be 'defined' in terms of 'gracet' and a predicate 'blacet', itself defined as follows:

(B) x is blacet if and only if *either* x has fewer than 99 facets and is blue *or* x has at least 99 facets and is green.

- (6) 'Gracet' and 'grue' are both constructed on the model.

(C) Gx iff $(Hx \ \& \ Jx) \vee (\sim Hx \ \& \ Kx)$, where J and K are logical contraries.

J. E. J. ALTHAM
Gonville and Caius College
Cambridge

REFERENCES

1. R. G. Swinburne, 'Grue', *Analysis*, 28 (1968), 123-8.

ANNOUNCEMENT

INTERNATIONAL KANT CONGRESS

The third INTERNATIONAL KANT CONGRESS will be held at the University of Rochester from 31 March to 3 April 1970. More detailed announcements will be made later, but inquiries and expressions of interest and intent are welcomed now. The committee on arrangements in the U.S.A. is L. W. Beck (*Rochester*) George Schrader (*Yale*) and John R. Silber (*Texas*); the European committee is Gottfried Martin (*Bonn*) and Ingeborg Heidemann (*Bonn*). Correspondence from the English-speaking world should be directed to Professor Beck; from Continental Europe and other parts of the world to Professor Martin.

Reviews

Wesley C. Salmon: *The Foundations of Scientific Inference* (Pittsburgh, University of Pittsburgh Press, 1966, \$2.50.) Pp. 157.

According to Professor Salmon, demonstrative inference is necessarily truth preserving inference. Inductive inference is legitimate inference from premises to acceptance of a conclusion which is non-demonstrative. The central problem discussed in this book is Hume's problem—i.e. the problem of justifying legitimate modes of non-demonstrative inferences.

For Salmon, the problem of induction so construed reduces to the problem of justifying inverse inference from observed frequencies to probability hypotheses which Salmon interprets as statements about limits of relative frequency in infinite sequences. Direct inference from a statistical hypothesis to the outcome of sampling is not, on his view, a mode of inference. If one knows that the chance of obtaining heads on a toss of a coin is 0.5, no conclusion can be reached legitimately regarding the outcome of the next toss or the relative frequency of heads on the next 100 tosses. 'Weights' can be assigned to such predictions for purposes of betting and practical policy making. But assigning weights involves no inductive inference (p. 94).

Nonetheless, Salmon does feel obliged to offer some justification for adopting certain procedures rather than others in assigning weights to outcomes of sampling even though this problem is not itself a problem of justifying a mode of inductive inference. Salmon argues that if a gambler were to judge the expected value of gambles offered to him using weights based on statistical hypotheses in the manner Salmon recognises as legitimate, then if the gambler were to have the opportunity of accepting gambles bearing constant expected value an infinite number of times, his 'net gain' in the long run would be equal to the expected value. This claim can be deduced from the statistical hypotheses on the basis of which weights are assigned when such hypotheses are understood as asserting limits of relative frequency in infinite sequences.

Thus, Salmon's argument depends critically upon his interpretation of probability in terms of limits of relative frequency in the infinitely long run. This interpretation is fraught with well-known difficulties not the least of which are those concerning the ontological status of infinitely long sequences. Salmon does consider alternative interpretations of probability of the logical and subjectivist varieties. He neglects, however, to discuss other interpretations of probability like those of Harald Cramer and R. B. Braithwaite which resemble his own in treating probability as an empirical or theoretical magnitude but which do not involve translation of probability statements into statements about relative frequency either in infinite or finite sequences.¹

¹ Braithwaite does allow frequency *models* of probability; but he denies that such models specify truth conditions for probability statements. Salmon does not seem to be clear on this point. See his footnote 107.

Salmon's failure to discuss such rival interpretations of probability is a serious omission. These interpretations have many of the merits which he claims for his own view. However, statistical hypotheses as interpreted by Cramer and Braithwaite do not deductively imply statements about the value of the limit of relative frequency. On their views, it is logically possible that the chance of a black chip being drawn from an urn is 0.5 and that every chip drawn in an infinite sample with replacement is black. The truth of the statistical hypothesis does not deductively guarantee that a gambler with an infinite amount of time will have a net gain 'in the long run' equal to the expected value determined by Salmon's method of weighting. By ignoring approaches like those of Cramer and Braithwaite, Salmon fails to take into account a serious challenge to his view that assignments of weights to outcomes of sampling can be justified without inductive inference from statistical hypotheses to outcomes of sampling.

But, as Salmon himself acknowledges, in the long run we are dead. Hence, assignment of weights ought to be justified by an appeal to the gains which will accrue to the gambler in that finite number of gambles in which he may follow Salmon's policy. In a footnote, Salmon admits that this problem must be faced and refers approvingly to papers where he discusses the short run. And, interestingly enough, in those papers he admits what he seems to deny in the book—to wit, that direct inference from statistical hypothesis to the outcome of finite sampling is a legitimate mode of inductive inference requiring justification. It would have been helpful had Salmon seen fit to clarify his position on this point.

Salmon attempts to show that other apparent modes of inductive inference (e.g. hypothetico-deductive reasoning) when rightly understood are mixtures of inverse inference, assigning weights on the basis of statistical assumptions and manipulations of the probability calculus. Hume's challenge reduces itself to the task of justifying inverse inference from observed relative frequencies in a finite segment of an infinite sequence to the limit of relative frequency in the sequence. Salmon does not pretend to have solved Hume's problem. However, he does believe that Reichenbach's programme for vindicating induction is on the right track.

A known-to-be-convergent rule *R* for inferring limits of relative frequencies from observed frequencies has the following property: if *R* is repeatedly applied to revise estimates of the limit in a sequence as the observed sample changes, it is demonstrable that if there is a limit the rule *R* will eventually identify it correctly.

According to Reichenbach and Salmon, if a sequence has a limit, a known-to-be-convergent rule is preferable to any alternative—as long as we have no factual information other than that contained in the observed sample. Moreover, a known-to-be-convergent rule is no worse than any alternative when there is no limit. So whether there is a limit or not, nothing is lost by adopting a known-to-be-convergent rule; and something might be gained.

Although this argument does not show why one of the infinitely many known-to-be-convergent rules is better than any other such rule, Salmon feels that it establishes the superiority of any such rule over rules which are not known-to-be-convergent. In this respect, Salmon feels that Reichenbach has made more progress in confronting Hume's problem than anyone since Hume.

This argument rests on the mistaken assumption that, when attempting to posit a limit of relative frequency, the only options open to us are methods of assigning values to such limits. But it is always open to us to deny the existence of a limit or to suspend judgment. If truth is preferable to error (and if Hume's problem is a problem this must be the case), then, when there is no limit, both of these options are better than known-to-be-convergent rules. Reichenbach's programme fails to get off the ground.

Salmon's book is written in a clear and elementary style. In the course of developing his argument, he offers many shrewd observations about alternative approaches to the problem of induction and the relations between frequentist, logical and personalistic interpretations of probability. The book has many of the earmarks of a text useful for beginners and advanced students alike. Beginners should be warned, however, not to take Reichenbach's myth about vindication too seriously.

ISAAC LEVI

Case Western Reserve University

Isaac Levi: *Gambling with Truth*. (New York: Alfred A. Knopf; London: Routledge and Kegan Paul, 1967. 63s.). Pp. xiv+246+v.

Professor Levi sets out to develop an account of rational belief on the analogy of practical decision theory. His starting point is the Bayesian principle that a rational decision-maker chooses, from the options open to him, that which maximises expected utility—summing, for each option, its utility on each possible outcome, weighted in accordance with the probability of that outcome. To make cognitive decisions analogous with this, Levi needs a cognitive counterpart to utility, for he refuses, rightly, to reduce belief to practice. He therefore introduces 'epistemic utility': the epistemic utility of accepting a particular answer to a question is a function both of its truth or falsehood and of its content, that is, the degree to which it gives a specific answer to the question. The investigator's aim is to 'replace agnosticism by true belief'—by 'agnosticism' Levi means ignorance, the state of having to suspend judgment—so that truth and 'relief from agnosticism' are the two desiderata which may have to be weighed against each other. One could guarantee 'truth' by accepting no more than the disjunction of all the alternative answers that were not deductively ruled out by the available evidence, but this 'skeptical' position would defeat one of the aims of science, which is to reach definite answers. Levi's central programme, then, which is completed less than half-way through the book, is to suggest how 'relief from agnosticism' and therefore content may be measured, to construct epistemic utility functions that balance truth with content, and to formulate an acceptance rule which tells us what answer is to be believed. The later chapters apply these suggestions to a variety of questions—e.g. to measuring degrees of belief, to Carnap's proposals for measuring confirmation, and to questions about direct and inverse statistical inference.

A key concept is that of an 'ultimate partition', that is, an exhaustive and exclusive set of alternative basic answers to a question—in a lottery problem, the

ultimate partition might be 'ticket 1 will win', 'ticket 2 will win', and so on, or it might be 'ticket 3 will win' and 'some ticket other than 3 will win'. Relief from agnosticism, and so content, is measured by giving the rejection of each basic alternative an equal degree of relief, so that the content of any (possibly disjunctive) answer H is m/n , where n is the number of basic alternatives not deductively ruled out by the evidence, and m the number of these that are excluded by H .

Another key concept is that of a ' q -index', i.e. an index of caution; q enters into the epistemic utility functions in such a way that as q approaches 0, the weight given to content decreases, and if $q = 0$ the investigator is committed to the above-mentioned 'skeptical' position, but as q approaches 1 the investigator becomes less cautious and more ruthless in rejecting alternatives with relatively low probabilities.

Thus an epistemic decision problem is relative (i) to an ultimate partition, (ii) to a body of evidence, with probabilities for the various answers on that evidence, and (iii) to a degree of caution. The proposed acceptance rule, Rule (A), is that we eliminate all basic alternatives that are inconsistent with the evidence, and also all those whose probabilities are less than q/n (with n as above), and then suspend judgment between all those not thus eliminated, i.e. accept as an answer the disjunction of all the surviving alternatives. This rule is backed up by a more elaborate but still fairly simple set of theorems about content, utility, etc., but since the construction of this theory involves several arbitrary moves, the justification of the theory as a whole must rest on the success of the acceptance procedure based on it in organising the cognitive decisions that we would actually be prepared to make.

One controversial feature is this relativisation to ultimate partitions, which means that what it is rational for me to believe depends partly on what question I ask myself. This has a bearing on 'deductive cogency', i.e. on the question whether rational belief in a set of statements commits one to believing all their deductive consequences. Thus if an urn contains 999,999 black balls and one white, and balls are drawn, and replaced, a billion times, Levi's system would lead one to accept, at each draw, the prediction that a black ball will be drawn, but also to accept the answer that a white ball will appear at least once in a billion draws (pp. 40-41). Levi resolves the inconsistency by saying that these are answers to different questions, relative to different ultimate partitions; against Kyburg he defends deductive cogency among those beliefs that are rational relative to any one ultimate partition. Here Levi's acceptance rule is superior to one that takes probability alone as a criterion of acceptance, for the latter is certain to clash with deductive cogency because the conjunction of two highly probable statements will be less probable. In Levi's system what requires rejection of a basic alternative is not low probability alone, but a probability that is low in relation to the number of elements in the partition, i.e. a low probability not compensated by high content. But one may still question the relativisation of deductive cogency to ultimate partitions; in a system of beliefs that were really rational, one would expect even the answers to different questions to cohere.

An associated controversial item is the measurement of the content of H simply by the proportion of basic alternatives that it excludes. Levi argues plausibly (pp. 163-5) that the content of H cannot be measured simply by the probability,

or degree of confirmation, of its denial. For this way of measuring content would ensure that when truth and content are weighed together the recommended policy would always be the sceptical one of accepting only the deductive consequences of the evidence and otherwise suspending judgment. But granted this, Levi's positive argument for his own measure (p. 166) is question-begging, and on the face of it different alternatives would seem able to have different content in the sense in which increasing content is an aim of science.

Rule (A) itself seems to have both good and bad consequences. Thus Levi gives an example where of four hypotheses only H_3 and H_4 are at first rejected, so that the recommended answer is $H_1 \vee H_2$; but then, since H_3 and H_4 have been rejected with a high degree of confidence, their exclusion can be included in the evidence, and when the problem is this restructured H_2 is rejected, leaving H_1 alone as the answer. As Levi admits, the strength of the conclusion reached has been increased by mere book-keeping (p. 151). This is a serious paradox. Levi argues that it is inescapable once it is allowed that an inductive conclusion can be added to the evidence, and he says, plausibly, that this should be allowed. But this conclusion is inescapable only given Levi's own kind of acceptance rule, and it is upon this that the paradox might cast doubt. Again, Levi shows (pp. 214-24) that Rule (A) yields the desired answer in direct statistical inference if the ultimate partition is into hypotheses about relative frequencies of different possible results, but not if it is into hypotheses that take account of the order in which different kinds of result occur; yet the latter would appear to represent an equally legitimate sort of question.

There are difficulties, then, in Levi's detailed proposals. And it is very hard to see how his method could be extended from the simple kinds of problem, involving clearly defined alternatives, to which it is most obviously applicable, to questions about the confirmation of genuine scientific hypotheses. Levi himself admits that this book only begins to sketch a theory; he repeatedly mentions lines of development that are still to be pursued. Nevertheless, his approach is extremely interesting, and his main suggestion that the aims of science point to a cognitive decision procedure that takes account of both probability and content is highly plausible.

Some minor features of the book are less admirable. Levi's style is often turgid; he cannot present philosophical issues clearly (e.g. when in chapter XIV he tries to outline a dispositional theory of statistical probability); his first and his last chapter display metaphysical cramps that seem to be due to 'behavioralist' influences; but one gets the broad impression that his heart is often in the right place. He brushes difficulties aside in a cavalier way; but often justifiably, since he wants to get on with a particular job; in the later chapters he touches, too briefly to contribute much of value, on many diverse issues. Here is one example of his slapdash manner. He accuses 'most philosophers who write in ethics' of subscribing to 'naïve cognitivism' (pp. 7-9), but an alternative to naïve cognitivism is the view (p. 11) that 'probability is the guide in life'. Levi seems unaware that this dictum is a quotation, slightly altered, from Bishop Butler's Introduction to *The Analogy of Religion*.

J. L. MACKIE
University College Oxford

Richard Schlegel: *Completeness in Science* (New York: Appleton-Century-Crofts, 1967, n.p.). Pp. xi+280.

It was fashionable a century ago to suggest that the important work of science had been accomplished since all that remained unfinished and unknown only required a mopping-up operation in order to determine natural constants with greater accuracy and to tidy up a few anomalies in radiation theory. In spite of its apparent naivety, the nineteenth-century question 'Is science finished?' has re-emerged in the twentieth century in a slightly different guise: 'Is science, or can it be, complete?' The work of Gödel and Church in mathematical logic as well as the qualms of Einstein, Podolsky and Rosen about the quantum mechanics have made the question of completeness again a fashionable one. What chiefly distinguishes contemporary discussions of this question from their Victorian predecessors is a considerably more sophisticated set of analytic tools as well as a radically different scientific context. But for all that, it is not clear that we can now give a more convincing answer to the question.

In this book, completeness serves as the peg on which the author hangs a number of themes, ranging from transfinite arithmetic to quantum theory and cosmology. His discussion of most of these issues is interesting and provocative, if not particularly original; but less satisfactory is what he has to say about completeness itself. He begins by distinguishing between descriptive and explanatory completeness, a distinction which he preserves in his final definition of completeness: 'a science is complete when it gives as much descriptive detail as is desired for the domain of the science and when the theoretical structure of the science satisfactorily explains all of the facts of the science' (p. 46). Besides the fact that Schlegel's definition makes completeness a psychological rather than an objectively characterizable feature of a theory, there are several puzzling things about the way he applies this notion of completeness to the physical science. I should like to draw special attention to two crucial aspects of his argument:

(a) Hoping to find some way in which the Gödel incompleteness theorem might bear on the empirical sciences (for otherwise his lengthy informal treatment of that theorem is irrelevant to the rest of the book), he argues that some scientific statements can be self-referential. The one example he gives (to which he refers throughout the book) is this: '... suppose that in Volume Z of a work of scientific description, the 11th sentence of Chapter 5 reads as follows: "The 11th sentence of Chapter 5 of Volume Z is not a valid sentence of scientific description"' (p. 80). Schlegel argues that this sentence is an undecidable one (provided our science is consistent), and suggests that this points to an important kind of incompleteness in science. One could reply that Gödelian sentences of the kind Schlegel contrives would never count as (well-formed) scientific descriptions; but even without rehearsing that familiar argument, one wants to say that it is only a highly trivial sense of incompleteness that Schlegel is discussing here and that it is not Gödelian completeness that one expects in science, however complete science might be in other respects. The basic shortcoming of Schlegel's analysis here is his failure to explore in sufficient depth the question of the *domain* of science. Although this notion occurs in his own *definiens* of

completeness and is clearly relevant to the question whether there are facts which science cannot explain, he prefers to pursue the pseudo-problems of self-referential scientific statements. Indeed, one suspects that Schlegel allows Gödelian statements in science precisely because he has not specified clearly enough what is the scientific domain.

(b) For a number of plausible reasons, Schlegel claims that quantum mechanics is incomplete. This is, of course, a wide-held view in the philosophy of physics and I shall not presume to question it here. What does seem very dubious, however, is Schlegel's claim that all future physical theories will share the incompleteness of quantum physics. Schlegel thinks it most unlikely that we shall ever find a theory which transcends the limitations imposed on description and prediction in quantum mechanics. His only argument to support the perpetuity of quantum incompleteness is the irrelevant one that 'physics in its development has tended not so much towards declaring an already established truth to be wrong but rather towards incorporating that truth into a broader theory' (p. 247). As an historical thesis, this may or may not be sound, but until we are shown that quantum incompleteness is 'an already established truth', Schlegel's argument for its permanence can scarcely be very convincing.

Not surprisingly, Schlegel's general conclusion is that science is essentially incomplete in both the descriptive and the explanatory sense. As he concedes, this is not exactly an unexpected conclusion, nor are his reasons for adopting it particularly original. He excuses himself by telling the reader that 'philosophy is an adumbration of the obvious' (p. 253) and that one should not expect startling conclusions from a philosophical treatment of science. One can only be grateful that some philosophers of science do not share Schlegel's conception of the subject and its content; from them, we might learn something about completeness in science!

L. LAUDAN

University College London

H. Regnéll: *Ancient Views of the Nature of Life: Three Studies in the Philosophies of the Atomists, Plato and Aristotle* (Lund: C. W. K. Gleerup, 1967, 35 Sw. Cr.). Pp. 267.

Of the three studies of the Greek Atomists which Regnéll presents in this book the first, which occupies three quarters of the book, examines Atomist views on the content and scope of the concept of life and compares them with those of their predecessors and contemporaries, especially Plato and Aristotle, while the second and third are concerned with the motion of atoms and with the nature of the gods in Atomist theory.

In the introduction to the first study Regnéll sets out a list of manifestations of life—such as bodily warmth, self-motion, self-preservation, the ability to reproduce, consciousness—which he regards as forming the material which provided the impulse for ancient theories on the nature of life and which accordingly

provide the articulation of his analysis of those theories. Regnéll shows in his chapter on the Atomists how attention to the manifestations of life shaped their general position according to which, (a) the scope of the concept of life excluded the world and heavenly bodies and plants (Epicurus), (b) the content of the concept involved a relation between body and soul of such a kind that the possibility of the independent existence of the soul was excluded. A feature of Atomist thought which Regnéll is concerned to emphasise is their view that all aggregates of atoms are continually gaining and losing individual atoms so that the permanence of any thing is to be seen as 'dynamic'. In his second chapter Regnéll examines the prominence of this idea in other Greek thinkers—Heraclitus, Empedocles, Plato, and Aristotle; although in the case of Aristotle it should be recognised that the type of identity manifested by the water of a river is regarded as an offshoot of specific identity and sharply distinguished from numerical identity (Topics A7). In the chapter of the *De Gen. et Corr.* (A5)—on the analysis of growth—to which Regnéll appeals at this point, Aristotle contrasts the permanence of form, which increases uniformly, with the constant change undergone by the matter in a growing organism. Since Aristotle identifies form with life, his account needs to be distinguished rather more sharply from that of the Atomists, who are certainly materialists to the extent of identifying a living thing with its matter. Generally, however, Regnéll's analysis of the similarities between the thinkers with which he deals is sound and illuminating.

The second and third studies arise out of two difficulties encountered in the first study. Although the Atomists agree with other thinkers in making the soul responsible for the motion of an organism, in what sense can self-motion be ascribed to soul atoms or indeed to any atoms? Regnéll's conclusion here is that (a) for Democritus no sense can be attached to a distinction between original and derived motion but the peculiar properties of soul atoms make them especially suitable for initiating lifelike movements, (b) for Epicurus atoms do have an original motion attributable to their weight and they also undergo the minute swerve which Epicurus introduced to avoid determinism—(Regnéll is no more successful than other commentators in explaining how the swerve and free will are connected). The final study argues that Epicurean theology is to be taken seriously and shows how the life enjoyed by the gods is conceptually compatible with more familiar forms of life.

Regnéll makes a good case against those who are so blinded by the opposition between, say, Plato and Epicurus on questions of grand metaphysics that they fail to appreciate the large area of common ground in their theories. This common ground consists not only in the information which they derive from experience, although primarily in this, but also in their conceptual apparatus. It is perhaps characteristic of the historian of science to concentrate on the similarities between these thinkers, while the historian of philosophy with equal justification insists on their opposition; for the opposition between Plato and Aristotle on the one hand and the Atomists on the other with regard to questions of the nature of the life and soul of man stems from their different views on the nature and function of philosophy. For Plato and Aristotle the value of intellectual life prompts an ontology of soul which can secure that value in reality, while the Atomists were led by their different values equally to provide a basis in fact for their ideals.

While there is little that is original in this book, it is a sensible and clearly argued statement of the present state of the problem. The documentation and bibliography are sound, and the style (it is translated from Swedish) is vigorous and clear.

J. D. G. EVANS
Sidney Sussex College,
Cambridge

S. J. Prokhovnik: *The Logic of Special Relativity*. (Cambridge University Press, 1967. 32s. 6d.; \$5.95.) Pp. xiv+128.

This book examines the various difficulties and controversies associated with the special theory of relativity. Since it was first formulated in 1905, special relativity has given rise to heated discussions among scientists and philosophers. To some it is the perfect theory—the very last word in physics; whereas to others it is full of logical contradictions. The author adopts a middle course, and wants to show that while the theory is ‘a monument to physical insight and mathematical logic . . . its interpretation is inevitably incomplete’.

The book is divided into six chapters and six appendices, one for each chapter. The appendices contain proofs, mainly mathematical, of the results used in the main chapters. This is an excellent arrangement, since one is then able to follow the main chain of arguments uninterrupted. This will also suit the less mathematically minded reader.

Chapters 1 and 2 give the general background of the special theory of relativity and the way Einstein arrived at it. The celebrated clock paradox is discussed. The author points out that although the controversy on this subject is generally described as Dingle versus ‘The Rest’, the latter are very much divided among themselves as to the correct interpretation of the time dilatation effect. He points out the different points of view adopted by the principal participants in the controversy. The different ways in which the paradox is ‘resolved’ are described in chapter 3. In chapter 4, the concept of ‘kinematic symmetry’ which originated with Milne, is discussed. This concept has been developed by several authors in different ways since Milne’s time. The author shows how it leads to the Lorentz transformation in one space+one time dimensions in an elegant way. However, as yet its extension to three space+one time dimensions has not been possible.

In chapter 5 the author examines the case for an aether. Although aether became unfashionable since the advent of special relativity, there are some who would like to return to it in view of the existence of controversy around special relativity. The author outlines an aether type theory which starts off with a preferred inertial frame I_0 , ‘relative to which the propagation of energy is isotropic and for which Newton’s laws of motion hold!’ A body moving relative to I_0 with velocity u is contracted in its length by a factor $\sqrt{(1-u^2/c^2)}$ along the direction of motion. He then describes the phenomena usually associated with special relativity in terms of these assumptions. The final chapter is devoted to cosmological considerations. The expansion of the universe does appear to single out a special local inertial frame (in which the expansion is isotropic).

The author describes a cosmological model in which special relativity—as reinterpreted in the previous Chapter—can be applied ‘in the large’.

The appendices describe useful results such as the invariance of Maxwell equations, derivation of Lorentz transformations, Möller’s discussion of the clock paradox, etc.

The book is well written and will be useful to the student of special relativity because it gives a concise description of the development of the subject from various angles. An extensive list of references to the works of others quoted or described in the text is given at the end of the book.

J. V. NARLIKAR
King’s College, Cambridge

The Correspondence of Isaac Newton, vol. iv: 1694-1709. (Ed. J. F. Scott) (C.U.P. for the Royal Society) 1967, 11 guineas. Pp. xxxvi+578.

The fourth volume of this great enterprise is the first under the editorship of Dr Scott who succeeded the late Professor Turnbull. A historian of mathematics of long standing and established reputation, Dr Scott has had no easy task in the current volume, for the period covered was probably the most varied of Newton’s career. The two principal themes are Newton’s relations with Flamsteed, the first Astronomer Royal, particularly in connection with data concerning the motion of the Moon, and Newton’s activities as Warden of the Mint from March 1696.

Unlike the previous volumes, this one contains few points of philosophical importance, but it covers much that is of psychological interest. Locke said that Newton was ‘a nice man to deal with, and a little apt to raise in himself suspicions where there is no ground’. Newton was anxious to get accurate data concerning the Moon in order to perfect his lunar theory and only Flamsteed could supply them. In a letter to Flamsteed of 17 November 1694 (Number 480), he dismissed the usual way of tackling the problem of the Moon’s motion as ‘bungling and tedious’. Instead, he claimed, ‘The method wch I propose to my self is first to get a general notion of the equations to be determined & then by accurate observations to determin them’. Newton soon became impatient with Flamsteed for what he regarded as dilatoriness in sending him the observations he needed. Neither Newton nor Flamsteed were masters of tact and their correspondence provides a vivid picture of how each succeeded in irritating the other beyond measure. One of the facts that increased Flamsteed’s ill-feeling was Newton’s friendship with Halley whom Flamsteed heartily detested for what he regarded as ‘his ill morals & abuseing religion’.

As Dr Scott points out, the letters in this volume show that Newton was not only a superlative genius but also a man of the highest principles. This is clearly revealed by his activities as Warden of the Mint. He took these so seriously that he almost ceased working in science. If he had wanted to, he could have treated the post as a well-paid sinecure. Instead, he was so active and conscientious in

stamping out currency offences and in dealing with other troubles that he set a much needed standard for integrity in the public service.

The present volume maintains the high level of this series and is graced with some attractive plates. We eagerly look forward to the remaining volumes.

G. J. WHITROW
Imperial College, London

Maurice Mandelbaum: *Philosophy, Science and Sense Perception*. (Baltimore; Johns Hopkins Press, 1967, 208.) Pp. x+262.

Three important points are made by Professor Maurice Mandelbaum in *Philosophy, Science and Sense Perception, Historical and Critical Studies*:

(1) That central to the philosophic and scientific work of Locke, Boyle and Newton is the view that inferences may be made from sense experience to the existence and nature of an independent physical reality.

(2) That the arguments used by phenomenologists, such as Hume, to prove that there can be no inference from experience to the existence of things existing independent of experience, themselves presuppose the existence of an independent physical reality.

(3) That the arguments of phenomenologists, such as Hume, are sufficient to refute a naive realism but insufficient to refute a critical realism which maintains that we can know some of the properties of physical objects as possessing any of the properties of objects as perceived by us.

Professor Mandelbaum's overall conclusion is that the findings of science are relevant to epistemology.

Of the three contentions, the first is the most incontrovertibly established. Professor Mandelbaum first states the textual sources for the widespread phenomenological interpretations of Locke, Boyle and Newton, and then shows that such passages do not require a phenomenological interpretation when placed within their full context. The passages that strongly suggest a commitment to a realistic atomism may then be accepted on face value. This task occupies the first two studies of *Philosophy, Science and Sense Perception*.

The last two studies, however, are far more controversial than the first two.

In the third study, Professor Mandelbaum seeks to show that the phenomenological arguments used by Hume and others presuppose the very doctrine that they are intended to refute: The existence of an independent physical reality. What Professor Mandelbaum does prove, it may be argued, is that these arguments presuppose that experience is objective. But it is not obvious that to assume the objectivity of experience is to assume the existence of an external world. It would have been of great interest if Professor Mandelbaum had dealt with the merits of a Kantian analysis of the objectivity of experience.

The main argument of the last study is that the arguments of phenomenologists such as Hume suffice to refute a naive realism, but are inadequate as a refutation of critical realism. According to Mandelbaum, the main difficulty in accepting critical realism is due to the discrepancy between the scientific account of physical objects and objects as they are perceived. This difficulty is alleviated when it is

realised that the methods employed by scientists to draw inferences from experience concerning the nature of physical reality are continuous with methods used in everyday life to draw inferences from experienced features of objects to as yet unexperienced but in principle experientiable features of objects. A key assumption of this last argument is that the procedures of common sense are correct when used to draw inferences from experience to the character of objects as they exist independent of all experience.

Regardless of one's assessment of the findings and arguments of the last two studies, *Philosophy, Science and Sense Perception* is an important statement of the case for a critical realism. And it demonstrates the relevance of studies in the history of philosophy to philosophy, if not the relevance of science to epistemology.

M. PRIMACK

The Illinois Institute of Technology

Allen G. Debus: *The Chemical Dream of the Renaissance* (Churchill College Overseas Fellowship Lecture No. 3, Heffer, Cambridge, 1968, 6s.). Pp. 40.

Dr Debus's lecture is a critical examination of the traditional view of the Scientific Revolution as 'the forward march of a new experimental method coupled with the powerful tool of mathematical abstraction'. Rejection of Aristotelian natural philosophy and the advocacy of a new science of nature, based on 'observation and experiment': these slogans were not the monopoly of the mechanical philosophers. The same objectives had been vociferously championed since the later sixteenth century by the Paracelsians and alchemists. Dr Debus suggests that their confidence in the imminent triumph of their own natural philosophy explains the preoccupation of Paracelsians with comprehensive university reform from the late sixteenth century. That is a valuable suggestion and supplements the social, political, and theological considerations which must be taken into account as the background to the educational endeavours of such men as Andreae, Alsted, Jungius, and Comenius. For them, it was chemistry (rather than the study of motion, or of mathematics—except in the sense of mystical Pythagoreanism) which was the true key of nature, 'to unlock the secrets of heaven and earth'.

The influence of these educational ideals is traced in England by Dr Debus, as illustrated by Robert Fludd and the would-be university reformers of the Puritan Revolution. As the mechanical philosophy came to be more systematically articulated, the struggle between the two approaches became sharper: Helmont, for example, considered the increasing dominance of mathematics in the study of nature a backward step, away from the truly experimental study of nature. Dr Debus's interesting and suggestive account casts greater light on the resistance of, say, Le Febvre and Thomas Willis to the corpuscular-mechanical approach, which they contrasted unfavourably with the 'ocular demonstrations' of chemistry—a theme which, in various guises, was to serve as the focus of controversy through the eighteenth and nineteenth centuries.

P. M. RATTANSI
*King's College,
Cambridge*

- Machine Intelligence 1*: N. L. Collins and D. Michie (eds.). (Edinburgh: Oliver & Boyd, 1967, 63s.). Pp. x+278.
- Machine Intelligence 2*: E. Dale and D. Michie (eds.). (Edinburgh: Oliver & Boyd, 1968, 70s.). Pp. x+252.
- Machine Intelligence 3*: D. Michie (ed.). (Edinburgh: University Press, 1968, 70s.). Pp. x+405.

These volumes are a useful source of information on current experimental and theoretical research into the problems of machine intelligence. They are the first of a series of reports from the *Machine Intelligence Workshop*, an annual symposium first held at Edinburgh University in 1965, which brings together many of the leading British workers on experimental programming, and covers a wide range of topics in the field. There are fifty-five papers in all, and I can refer to only a few of them—but the range of problems discussed can be roughly classified as follows: the abstract foundations of programming (including such questions as how to prove two programmes equivalent); mechanised mathematics and theorem proving; machine learning and heuristic programming; cognitive processes; pattern recognition; the development of man-machine interaction and problem-oriented languages.

The main emphasis of these papers is on the discussion and practical realisation of *artificial intelligence*, rather than on the *computer simulation* of cognitive processes. That is, the primary aim is to develop machine-programmes which will in fact solve certain classes of problems, and the question whether the method of solution is analogous to that used by human problem solvers is secondary. But the two approaches, while they can be distinguished, cannot be divorced; for it is often found that the most efficient way of producing an artificial intelligence is to opt for some degree of mimicry of the strategies and heuristics involved in natural intelligence. For example, a radically new type of programme for the syntactic analysis of English sentences results from the decision to try to incorporate the methods used by the human speaker in understanding sentences; this approach produces a powerful programme which is much more efficient than its predecessors. Thus the computer simulation of psychological processes may be motivated not only by the theoretical aim of modelling some specific psychological theory, but also by the pragmatic aim of finding some mechanised way of solving a practical problem. Many different problems will raise the same basic principles, such as how to programme the abilities to generalise from incomplete data, to abstract from large amounts of data, and to adapt to the constraints of new environments, benefitting from mistakes and successes. The more such abilities can be programmed, the more flexible and 'intelligent' will be the machine's performance, and the less restricted to problems of a specific type.

Among the many papers on machine learning and heuristic programming those by Amarel and Doran are particularly interesting. Amarel (III: 10) discusses the effects of problem-formulation on the efficiency of problem-solving. He concentrates on problems about actions, where we can specify an initial situation, a terminal situation, a set of feasible actions and a set of constraints that restrict the applicability of actions. He takes the *Missionaries and Cannibals* problem and shows how successive changes in formulation each increase the

ease of solution. Elementary actions may be combined into macro-actions, redundant conditions eliminated, symmetries in the problem-space exploited, and critical points identified which form a higher-level subspace within which the problem may be schematically solved, the detailed solution being generated later. The choice of appropriate basic elements and attributes, and of appropriate representations of action-rules and problem-states, is crucial. This choice is less easy to mechanise than is the recognition of useful properties (e.g. symmetries, redundancies) of the problem-space, given some appropriate representation. Doran (III: 13) describes an automaton which combines 'insight' and trial-and-error to find its way home, its path tending to improve on successive trials but never becoming stereotyped. Doran's interest here is to investigate the relation of the organism to its environment, and to stress the importance of the classification of possible environments from the standpoint of the organism. The 'subjective' environment of the automaton is crucial in determining its performance, and allows for the analogue of insight in behaviour. Variable parameters determining the relative costs of exploration and exploitation (i.e. searching for some better plan or relying upon the best plan reached so far) affect the overall pattern of behaviour: these parameters include size of memory-store, rigour of evaluation-standards, depth of lookahead in planning, and the 'thinking' time used in planning. Changes in these parameters affect the time-cost of decision-making, with resultant changes in performance. Doran shows that considerations such as these must enter into any computer simulation of instinctive behaviour, and sketches an automaton with two basic goals ('warmth' and 'food') which may conflict with one another in certain environments.

'Will seeing machines have illusions?' asks Gregory, and concludes that they will (I: 11). He argues that the errors involved in visual information processing are largely due to the logical difficulties involved in seeing the world, difficulties which we should expect to encounter in any system set to solve the problem. The problem, basically, is that of developing efficient strategies for extrapolation beyond the immediate sensory evidence, so that minimal visual cues can be relied on in the prediction of future events and the production of appropriate action. The correlation of visual and tactual cues (associated in the context of active movement among the objects of the environment) provides for the building up of interpretative hypotheses linking the two-dimensional images with the three-dimensional objects they represent. The images are in themselves infinitely ambiguous, and if the environment changes so that the hypotheses already learnt are no longer appropriate, illusions will result. Logically paradoxical illusions may arise if the visual system involves parallel processing, for the information simultaneously relayed by two channels may conflict—but these illusions are relatively easy to detect. The difficulties of manoeuvring objects in outer space—where many normal depth-cues are missing—include the problem of the three-dimensional interpretation of visual images, and illusion may lead to disastrous mistakes. The more adaptable we had made our seeing machines, the less readily we should be able to understand the (two-dimensional) information they might relay to us from an alien world, for their interpretative hypotheses would have changed but ours would not. Gregory suggests that we might find ourselves in the position of the blind man, unable ultimately to understand the man or machine that had successfully learnt to see the alien world. He does not discuss

in detail what sorts of basic difference might (logically) exist, and what sorts of illusion and misunderstanding might therefore arise.

The complex problems involved in building a speech recognition machine are indicated by Hill's paper in vol. I. Speech recognition by machine has not yet been achieved, even for the class of digits-spoken-by-random-passers-by, despite the large amount of effort which has been devoted to the problem. The reason is basically that there are no invariant acoustic data groupings which represent what the linguist classifies as 'phonemes', and there is often not sufficient information in the acoustic signal alone to determine the word being spoken. Psycholinguistic experiments have shown that a man who recognises a word 'by ear' is in fact relying upon many other levels of information, including his knowledge of the speaker's individual characteristics, of speech statistics, of syntactical rules and of semantics. Until such features can be built into a machine which is basically an acoustic recogniser, we cannot expect a machine to recognise speech successfully. Speech statistics should not be too difficult to incorporate, but what of higher level skills such as knowledge of syntactical rules? Here we may turn to two interesting papers by Thorne and his co-workers, on the machine analysis of syntactic structure (II: 11 and III: 18).

These papers describe a programme which differs from most existing automatic parsers in that it does not require a complete dictionary of the language under consideration, and so is capable of analysing an infinite number of sentences. Moreover, the programme simulates the way in which the English speaker perceives the syntactic structure. Thus it assigns deep and surface structure analyses simultaneously, passing through the sentence once only, and producing alternative possible analyses at the end of this single pass in those cases where the sentence is syntactically ambiguous. The programme assigns syntactic labels to an infinite number of words, and can deal with 'All brillig were the slithy toves' much as the human does. We say, not 'I know *tove* is a noun, therefore the syntactic structure is such-and-such', but 'The syntactic structure is apparent, therefore *tove* is a noun'; and the programme, too, first analyses the syntactic structure and then concludes that *toves* must be a plural noun. The programme cannot dispense with a dictionary completely, as certain words play a major role in determining the syntax. Such 'closed class' words (e.g. determiners, prepositions, pronouns), together with certain inflectional and derivational endings and auxiliary verbs, are included in a finite dictionary, and they determine the syntactic structure. The part that 'open class' words are playing can then be assessed from this structure. Also included in the dictionary are the exceptions to any mechanical routine for identifying inflected endings (e.g. *lens*, *hatred*), and certain special verbs like *make*, *prove* and *find*, which can give rise to syntactic ambiguities (He made a good dinner for the cannibal). The closed-class dictionary requires at most 2,000 items. The programme also requires a representation of a (transformational) grammar supplying rules for sentence structure, such as that adjectives do not normally follow nouns. The operation of the analyser is a process of making and checking predictions about syntactic structure. It does not first try one analysis and then backtrack to see if any others are possible; instead, it develops all possible analyses simultaneously, relying on its 'memory' of decisions made up to the current point. This makes it much more efficient than most existing parsers, which give information about surface structure only.

Constructions involving *and* and *or* cause some difficulty, for we are hazy about the grammar of these terms; the grammar of *as* and *than* being even less clear, the programme is totally unable to handle them, and mechanisation must wait upon a more satisfactory grammatical analysis. One result of not using a full English dictionary is that an incorrect analysis is sometimes produced as well as the correct one (e.g. 'The cat adores fish' is treated as ambiguous, like 'The girl guides fish'). To eliminate most of these incorrect analyses the programme would require not only form-class information, but the information that certain nouns are abstract or concrete, verbs transitive or intransitive, taking only animate objects, and so on. The programme might build up a dictionary providing such information, given a few basic items as starting-points; but what these items should be is an interesting and unanswered question—should they be verbs or nouns, and which ones, and how many? The method of dictionary-building suggested would outlaw such expressions as 'He surprised Whitehall', on the grounds that *surprise* takes only animate objects, but perhaps it is reasonable to ignore metaphor until the more straightforward sentences can be coped with. If a syntactic analyser of this type, incorporating semantic information such as that discussed, were to be added to Hill's basic acoustic recogniser the prospects for a speech recognition machine would be much improved.

I have said a little about some of the papers I found most interesting, but my choice will not have reflected the concerns of all readers of this review. Those interested in machine intelligence should turn to these volumes (and their future successors in the series) for further details of current work on the field.

MARGARET A. BODEN
University of Sussex

Recent Publications

RECENT PUBLICATIONS ON THE PHILOSOPHY OF SCIENCE

(a) BOOKS RECEIVED FOR REVIEW

- William P. Alston: *Philosophy of Language*. (London: Prentice-Hall International. 1968. 46s. 6d.) pp. xiv+113
- M. R. Ayers: *The Refutation of Determinism*. (London: Methuen. 1968. 37s. 6d.) pp. viii+188
- Ian G. Barbour: *Issues in Science and Religion*. (London: S.C.M. Press. 1968. 30s.) pp. x+470
- L. W. Beck and R. L. Holmes: *Philosophic Enquiry: An Introduction to Philosophy*. (London: Prentice-Hall International. 1968. 65s.) pp. xii+448
- Saul Benison: *Tom Rivers: Reflections on a Life in Medicine and Science. An Oral History Memoir*. (M.I.T. Press. 1967. \$17.50) pp. xxi+682
- H. Bondi: *Assumption and Myth in Physical Theory*. (Cambridge University Press. 1967. 17s. 6d.) pp. v+88
- R. B. Braithwaite: *Scientific Explanation*. (Cambridge University Press. 1968. (1st paperback edn) 15s.) pp. xii+376
- Jean Ford Brennan: *The Elegant Solution*. (London: D. van Nostrand Co. 1967. 48s.) pp. xiii+170
- May Brodbeck, Ed.: *Readings in the Philosophy of Social Sciences*. (London: Collier-Macmillan, 1968. 95s.) pp. 640
- Constantin Brunner: *Science, Spirit, Superstition: A New Enquiry into Human Thought*. Abridged and translated by Abraham Suhl; revised and edited by Walter Bernard. (London: George Allen & Unwin. 1968. 80s.) pp. 585
- M. Bunge, Ed.: *Studies in the Foundations, Methodology, and Philosophy of Science. Vol. 2: Quantum Theory and Reality*. (Berlin-Heidelberg-New York: Springer-Verlag. 1967. DM 29.60, \$7.40) pp. vi+117
- Huntington Cairns: *Legal Philosophy from Plato to Hegel*. (Oxford University Press. 1967. 32s.) pp. xv+583
- Chao Yuen Ren: *Language and Symbolic Systems*. (Cambridge University Press, 1968. 30s.) pp. xv+240
- Jean Charon: *Man in Search of Himself: How the Intellectual Caveman can Break Free of his Den*. (London: George Allen & Unwin, 1967. 35s.) pp. 210
- Percy S. Cohen: *Modern Social Theory*. (London: Heinemann. 1968. 35s.) pp. xi+247
- N. L. Collins and D. Michie: *Machine Intelligence* (3 vols). (Edinburgh: Oliver & Boyd. 1967. 63s. each vol.) pp. x+278
- A. Crosby: *Creativity and Performance in Industrial Organization*. (London: Tavistock Publications. 1968. 42s.) pp. xi+195
- J. N. Crossley, Ed.: *Sets, Models and Recursion Theory*. (Amsterdam: North-Holland Publishing Co. 1967. 50s.) pp. 331
- Frederick J. Crosson, Ed.: *Science and Contemporary Society*. (London: University of Notre Dame Press. 1968. 75s. 6d.) pp. xxv+251
- George H. Daniels: *American Science in the Age of Jackson*. (New York: Columbia University Press. 1968. \$7.95) pp. 282
- Allen G. Debus: *The Chemical Dream of the Renaissance*. (Churchill College Fellowship Lecture no. 3. Cambridge: W. Heffer & Sons. 1968. 6s.) pp. 40
- F. P. Dickson: *The Bowl of Night*. (London: Macmillan. 1968. n.p.) pp. xii+228
- Frank B. Ebersole: *Things We Know: Fourteen Essays on Problems of Knowledge*. (University of Oregon Books. 1968. \$7.50) pp. viii+304
- Paul Edwards, Ed. in Chief: *The Encyclopaedia of Philosophy* (8 vols). (London: Collier-Macmillan, 1967. n.p.)

- W. M. Elsasser: *Atom and Organism. A New Approach to Theoretical Biology*. (Oxford University Press. 1966. 36s.) pp. viii+143
- Marvin Farber: *The Foundation of Phenomenology*. (State University of New York Press. 1943. n.p.) pp. xi+585
- W. Farnsworth Loomis: *The God B Within*. (New York: October House. 1968. \$5.95) pp. ix+117
- Herbert Feigl: *The 'Mental' and the 'Physical': The Essay and a Postscript*. (University of Minnesota Press. 1968. 43s.) pp. vii+177
- A. Fischer: *Die philosophischen Grundlagen der wissenschaftlichen Erkenntnis*. (Vienna: Springer-Verlag. 1967. n.p.) pp. vi+226
- Richard M. Gale: *The Language of Time*. (London: Routledge & Kegan Paul, International Library of Philosophy and Scientific Method. 1968. 40s.) pp. viii+248
- Johan Galtung: *Theory and Methods of Social Research*. (London: George Allen & Unwin. 1967. 70s.) pp. 534
- D. G. Garon: *Relativity for Psychology: A Causal Law for the Modern Alchemy*. (New York: Philosophical Library. 1968. \$7.95) pp. 338
- T. Gold, Ed.: *The Nature of Time*. (Cornell University Press. 1967. \$8.75) pp. ix+224
- Rubin Gotesky: *Personality: The Need for Liberty and Rights*. (New York: Libra Publishers Inc. 1967. \$3.50) pp. 92
- Roger Hahn: *Laplace as a Newtonian Scientist*. (Los Angeles: William Andrews Clark Memorial Library, University of California. 1967. n.p.) pp. 23
- J. B. S. Haldane: *Science and Life. Essays of a Rationalist*. (London: Pemberton Publishing Co., The Humanist Library. 1968. 15s.) pp. x+213
- Pieter J. van Heerden: *The Foundations of Empirical Knowledge*. (The Netherlands: N. V. Uitgeverij Wistik-Wassenaar. 1968. n.p.) pp. xii+143
- Pieter J. van Heerden: *The Foundations of Physics*. (The Netherlands: N. V. Uitgeverij Wistik-Wassenaar. 1967. n.p.) pp. x+100+4
- Jacques Herbrand: *Ecrits Logiques*. (Presses Universitaires de France. 1968. n.p.) pp. 244
- R. J. Hirst: *Philosophy: An Outline for the Intending Student*. (London: Routledge & Kegan Paul. 1968. 25s.) pp. vii+167
- Willard C. Humphreys: *Anomalies and Scientific Theories*. (New York: Freeman, Cooper & Co. 1968. \$5.00) pp. 318
- Nikolai Iribadjakov: *Philosophia i Biologia*. (Sofia: Nauka i Izskuctvo. 1967. n.p.) pp. 495
- C. G. Jung: *Collected Works*. Vol. 13. *Alchemical Studies*. Ed. Herbert Reed, Michael Fordham, Gerhard Adler. Trans. R. F. C. Hull. (London: Routledge & Kegan Paul. 1968. 63s.) pp. xiv+453
- J. J. Kockelmans: *Philosophy of Science: The Historical Background*. (New York: Collier-Macmillan. 1968. 84s.) pp. xiii+496
- Arnold Koslow, Ed.: *The Changeless Order: The Physics of Space, Time and Motion*. (New York: George Braziller. 1967. \$7.59) pp. viii+328
- A. Koyré: *Metaphysics and Measurement: Essays in the Scientific Revolution*. (London: Chapman & Hall. 1968. 42s.) pp. x+165
- G. Kreisel and J. L. Krevine: *Elements of Mathematical Logic*. (Amsterdam: North-Holland Publishing Co. 1967. 76s.) pp. vii+221
- Imre Lakatos, Ed.: *Problems in the Philosophy of Mathematics*. (Amsterdam: North-Holland Publishing Co. 1967. 84s.) pp. xv+241
- Imre Lakatos, Ed.: *The Problem of Inductive Logic*. (Amsterdam: North-Holland Publishing Co. 1968. 105s.) pp. viii+417
- Imre Lakatos and Alan Musgrave, Eds: *Problems in the Philosophy of Science*. (Amsterdam: North-Holland Publishing Co. 1968. 128s.) pp. ix+448
- Susanne K. Langer: *Mind: An Essay on Human Feeling*. Vol. 1. (Oxford University Press. 1967. 80s.) pp. xxiii+487
- Henry M. Leicester, Ed.: *Chymia 12: Annual Studies in the History of Chemistry*. (University of Pennsylvania Press. 1967. \$8.50) pp. 236
- Emilio Roig de Leuchsenring: *Medicos y Medicina en Cuba: Historia, Biografia, Costumbrismo*. (Museo Histórico de las Ciencias Médicas, La Habana, Cuba. 1965. n.p.) pp. 269
- Martin Levey and Noury Al-Khaledy: *The Medical Formulary of Al-Samarquandi*. (University of Pennsylvania Press. 1967. \$15.00) pp. 382

- G. O. Lindholm: *Fehlerquellen der Sog. ASW-Versuche*. (Helsinki: Suomalainen Tiedakatemia. 1967. n.p.) pp. 230
- R. B. Lindsay: *The Nature of Physics: A Physicist's View of the History and Philosophy of his Science*. (Rhode Island: Brown University Press. 1968. \$7.50) pp. vii+212
- R. H. Lineback: *The Philosopher's Index*, Vol. 1, No. 4. (Ohio: Bowling Green University. 1967. \$2.50) pp. ix+66
- John Lyons: *Introduction to Theoretical Linguistics*. (Cambridge University Press. 1968. 70s.) pp. x+519
- Donald M. Mackay: *Freedom of Action in a Mechanistic Universe*. (Cambridge University Press. 1967. 5s.) pp. 39
- Edward H. Madden, Rollo Handy and Marvin Farber, Eds: *Philosophical Perspectives On Punishment*. (Illinois: Charles C. Thomas. 1968. \$6.75) pp. viii+150
- E. J. Miller and A. K. Rice: *Systems of Organization: The Control of Task and Sentient Boundaries*. (London: Tavistock Publications. 1967. 50s.) pp. xviii+286
- R. Näätänen: *Selective Attention and Evoked Potentials*. (Helsinki: Suomalainen Tiedakatemia. 1967. n.p.) pp. 225
- Joseph Needham: *Order and Life*. (M.I.T. Press. 1968. 18s.) pp. xvii+175
- Kelvin van Nuys: *Is Reality Meaningful?* (New York: Philosophical Library. 1966. \$10.00) pp. xiii+610
- Karl Pearson: *Early Statistical Papers*. (Cambridge University Press. 1967. 25s.) pp. 557
- Arthur N. Prior: *Papers on Time and Tense*. (Oxford: Clarendon Press. 1968. 30s.) pp. viii+166
- Vincent S. Potter: *Charles Peirce on Norms and Ideals*. (University of Massachusetts Press. 1967. \$7.50) pp. xiii+224
- Georges Poulet: *The Metamorphoses of the Circle*. (Oxford University Press. 1967. 80s.) pp. xxvii+400
- Winwood Reade: *The Martyrdom of Man*. (London: The Humanist Library, Pemberton Publishing Co. 1968. 15s.) pp. xvi+444
- H. Regnéll: *Ancient Views on the Nature of Life: Three Studies in the Philosophies of the Atomists, Plato and Aristotle*. (Lund: C. W. K. Gleerup. 1967. 35 Sw. Cr.) pp. 267
- J. L. Rigal, Ed.: *Le Temps et la Pensée Physique Contemporaine*. (France: Dunod. Réalités de la Science 4. 1968. n.p.) pp. 150
- Lynn E. Rose: *Aristotle's Syllogistic*. (Illinois: Charles C. Thomas. 1968. \$8.50) pp. vii+149
- Paolo Rossi: *Francis Bacon. From Magic to Science*. Trans. S. Rabinovitch. (University of Chicago Press. 1968. \$5.95) pp. xvii+280
- Dagobert D. Runes: *Treasury of Thought*. (New York: Philosophical Library. 1966. \$6.00) pp. 395
- David D. Rutstein: *The Coming Revolution in Medicine*. (M.I.T. Press. 1967. \$4.95) pp. ix+180
- Jan Rzewuski: *Field Theory. Vol. I: Classical Theory*. (London: Iliffe Books. 1967. (1st British edn.) 55s.) pp. 303
- H. Sachsse: *Die Erkenntnis des Lebendigen*. (Brunswick: Friedr. Vieweg & Sohn. 1968. n.p.) pp. 283
- Wesley C. Salmon: *Logic*. (London: Prentice-Hall International. 1968. 46s. 6d.) pp. xiv+113
- Israel Scheffler: *Science and Subjectivity*. (New York: Bobbs-Merrill Co. 1967. \$2.25) pp. vi+132
- F. O. Schmitt, T. Melnychuk, G. C. Quarton and G. Adelman, Eds: *Neurosciences Research Symposium Summaries. Vol. II*. (M.I.T. Press. 1967. \$12.00) pp. xiii+642
- I. M. Sechenov: *Los Reflejos del Cerebro*. (Museo Histórico de las Ciencias Médicas, La Habana, Cuba. 1965. n.p.) pp. 231
- Wilfred Sellars: *Philosophical Perspectives*. (Illinois: Charles C. Thomas. 1967. \$14.75) pp. vii+424
- Wilfrid Sellars: *Science and Metaphysics*. (London: Routledge & Kegan Paul. 1968. 35s.) pp. x+246
- Theo. G. Sinnige: *Matter and Infinity in the Presocratic Schools and Plato*. (The Netherlands: Vangorcum Assen. 1968. Hfl 24.50) pp. 252

278 Recent Publications

- L. Sodomka: *Structure and Properties of Solids*. Trans. K. A. Hay. (London: Iliffe Books. 1967. 15s.) pp. 175
- R. J. Strutt (Fourth Baron Rayleigh): *Life of John William Strutt* (Third Baron Rayleigh). (University of Wisconsin Press. 1968. \$10.00) pp. xxii+433
- C. W. Valentine: *The Experimental Psychology of Beauty*. (London: Methuen. 1968. 25s.) pp. 434
- Hermann Vetter: *Wahrscheinlichkeit und logischer Spielraum*. (Tübingen: J. C. B. Mohr (Paul Siebeck). 1967. DM 19.80) pp. viii+115
- Z. Vintř: *Transmutations of Nuclei*. (London: Iliffe Books. Physics Paperbacks 8. Ed. P. S. Allen. 1968. 15s.) pp. 118
- F. Waismann: *How I See Philosophy*. Ed. R. Harré. (London: Macmillan. 1968. 63s.) pp. viii+260
- F. Waismann: *Ludwig Wittgenstein und der Wiener Kreis*. Ed. B. F. McGuinness. (Oxford: Blackwell. 1968. 45s.) pp. 265
- W. H. Watson: *Understanding Physics Today*. (Cambridge University Press. 1967. 10s. 6d.) pp. xiii+218
- D. T. Whiteside, Ed.: *The Mathematical Papers of Isaac Newton. Vol. II. 1667-1670*. (Cambridge University Press. 1968. £10 10s.) pp. xxii+520
- L. L. Whyte: *Internal Factors in Evolution*. (London: Social Science Paperbacks & Tavistock Publications. 1968. 12s. 6d.) pp. xxiii+81
- D. Wiggins: *Identity and Spatio-Temporal Continuity*. (Oxford: Blackwell. 1967. 14s.) pp. viii+83
- Roger J. Williams: *You Are Extraordinary*. (New York: Random House. 1967. \$5.95) pp. x+252
- John Ziman: *Public Knowledge: The Social Dimension of Science*. (Cambridge University Press. 1968. 22s. 6d.) pp. xii+154
- Peter Zollinger: *The Political Creature: An Evolutionary Reorientation*. (New York: George Braziller. 1967. \$7.50), pp. vi+303

(b) ARTICLES

General

- P. Apostol: Actual trends in logic and methodology of science. *Progresele Stiintei* 3 (1967), 339-46
- S. G. Brush: Mach and Atomism. *Synthese* 18 (1968), 192-215
- Mario Bunge: Technology as applied science. *Technol. Cult.* 7 (1966), 329-47
- I. Hacking: A language without particulars. *Mind* 77 (1968), 168-85
- A. Joltea: Modern cosmological thinking. *Progresele Stiintei* 3 (1967), 408-16
- B. M. Kedrov: On the dialectics of scientific discovery. *Vopr. Filosofii* 20 (1966), 26-37
- P. D. Shaw: On the logic of particularity-assumptions. *Mind* 77 (1968), 186-205
- V. de Sousa Alves: A logica de Leibniz. *Rev. Port. Filosof.* 23 (1967), 425-46
- James A. Snedden: Historical estimation. *Phil. Phenom. Res.* 28 (1967), 241-51
- M. Strauss: Einstein's theories and the critics of Newton: An essay in logico-historical analysis. *Synthese* 18 (1968), 251-84
- W. Yourgrau and A. van der Merwe: Did Ernst Mach 'miss the target'? *Synthese* 18 (1968), 234-50

Epistemology, Semantics, General Philosophical

- Laird Addis: Particulars and acquaintance. *Phil. Sci.* 34 (1967), 251-9
- J. C. D'Alessio: On subjunctive conditionals. *J. Phil.* 64 (1967), 306-9
- M. Bunge: Do operational definitions of physical concepts exist? *Vopr. Filosofii* 20 (1966), 66-74
- R. S. Cohen: Ernst Mach: physics, perception and the philosophy of science. *Synthese* 18 (1968), 132-70
- Charles B. Daniels: Immediate knowledge: Ayer, Strawson and Shoemaker. *Theoria* 33 (1967), 176-88
- Donald Davidson and Zeno Vendler: Symposium: causal relations. *J. Phil.* 21 (1967), 691-703, 704-13
- F. I. Dretske: Can events move? *Mind* 76 (1967), 479-92

- Helena Eilstein: Some moral problems of science. *Stud. Filozof.* 3 (1967), 73-106
- Carlo Giannoni: Quine, Grünbaum and the Duhemian Thesis. *Notis* 1 (1967), 283-97
- Mary Hesse: Fine's criteria for meaning change. *J. Phil.* 65 (1968), 46-52
- Herbert Hochberg: Dispositional problems. *Phil. Sci.* 34 (1967), 10-17
- Jerzy Kmita: Scientific explanation and metaphor. *Stud. Filozof.* 3 (1967), 143-60
- Michael E. Levin: Fine, mathematics and theory change. *J. Phil.* 65 (1968), 52-6
- Erwin Levy: On the problem of the possibility of a perceptual world-in-common. *Phil. Phenom. Res.* 28 (1967), 48-57
- Norman Malcolm: The conceivability of mechanism. *Phil. Rev.* 67 (1968), 45-72
- William E. Mann: Definite descriptions and the ontological argument. *Theoria* 33 (1967), 211-29
- Henry Margenau: Symposium: quantum mechanics, free will and determinism. *J. Phil.* 21 (1967), 714-24
- Paul E. Meehl: Theory testing in psychology and physics: a methodological paradox. *Phil. Sci.* 34 (1967), 103-15
- Alexander P. D. Momelatos: Aristotle's 'powers' and modern empiricism. *Ratio* 9 (1967), 97-104
- Douglas Odegard: Arthur Pap and the paradox of analysis. *Theoria* 33 (1967), 230-45
- Patricia Sanborn: Objectification and self-knowledge: a critical examination. *Phil. Phenom. Res.* 28 (1967), 39-47
- Kenneth F. Schaffner: Approaches to reduction. *Phil. Sci.* 34 (1967), 137-47
- G. Schlesinger: Confirmability and determinism. *Phil. Q.* 18 (1968), 29-39
- Robert K. Shope: Meaning. *Inquiry* 10 (1967), 421-38
- Frederick A. Siegler: Unconscious intentions. *Inquiry* 10 (1967), 251-67
- Morton White: On what could have happened. *Phil. Rev.* 67 (1968), 73-89

Logic, Probability

- Edwin B. Allaire: Things, relations and identity. *Phil. Sci.* 34 (1967), 260-72
- H. G. Bohnert: Communication by Ramsey—sentence clause. *Phil. Sci.* 34 (1967), 341-7
- E. Bolker: A simultaneous axiomatisation of utility and subjective probability. *Phil. Sci.* 34 (1967), 333-40
- Hugh S. Chandler: Excluded middle. *J. Phil.* 64 (1967), 807-14
- Chung-Ying Cheng: Requirements for the validity of induction: an examination of Charles Peirce's theory. *Phil. Phenom. Res.* 28 (1968), 392-402
- K. M. Colby and H. Enea: Inductive inference by intelligent machines. *Scientia* 103 (1968), 43-52
- Sven Danielsson: Modal logic based on probability theory. *Theoria* 33 (1967), 189-97
- I. Hacking: Slightly more realistic personal probability. *Phil. Sci.* 34 (1967), 311-25
- G. Harman: Detachment, probability and maximum likelihood. *Notis* 1 (1967), 401-11
- Ono Katsuji: Does mathematics need something other than logic? *Ann. Ass. Phil. Sci.* 3 (1968), 7-18
- C. Onicescu: Present-day problems of the theory of probability. *Progresele Stiintei* 3 (1967), 445-9
- Dag Prawitz: Completeness and Hauptsatz for second order logic. *Theoria* 33 (1967), 246-58
- Wesley C. Salmon: Symposium: the philosophy of Rudolf Carnap—Carnap's inductive logic. *J. Phil.* 21 (1967), 725-38
- Leonard J. Savage: Difficulties in the theory of personal probabilities. *Phil. Sci.* 34 (1967), 305-10
- Leonard J. Savage: Implications of personal probability for induction. *J. Phil.* 64 (1967), 593-607
- A. Shimony: Amplifying personal probability. *Phil. Sci.* 34 (1967), 326-32

Exact sciences

- Mario Bunge: Physical axiomatics. *Rev. Mod. Phys.* 39 (1967), 463-74
- Mario Bunge: Mach's critique of Newtonian mechanics. *Am. J. Phys.* 7 (1966), 585-96
- P. Cartianu: From axiomatisation in mathematics to axiomatisation in physics. *Progresele Stiintei* 3 (1967), 321-7
- John Earman: On going backwards in time. *Phil. Sci.* 34 (1967), 211-22

- John Earman: Irreversibility and temporal asymmetry. *J. Phil.* **64** (1967), 543-9
- Brian Ellis and Peter Bowman: Conventionality in distant simultaneity. *Phil. Sci.* **34** (1967), 116-36
- Adolf Grünbaum: Are 'infinity machines' paradoxical? *Science* **159** (1968), 396-406
- Adolf Grünbaum: Reply to Hilary Putnam's 'An examination of Grünbaum's philosophy of geometry'. In R. S. Cohen and M. W. Wartofsky (Eds), *Boston Studies in the Philosophy of Science*. (Dordrecht: D. Reidel, 1968.) 1-150
- Adolf Grünbaum: 'The anisotropy of time. In T. Gold and D. L. Schumacher (Eds), *The Nature of Time*. (Cornell University Press, 1967), 149-86
- Martin Hollis: Times and spaces. *Mind* **76** (1967), 524-36
- G. Holton: Influences on Einstein's work in relativity theory. *Am. Schol.* **37** (1967), 59-79
- A. Koslow: Mach's concept of mass: program and definition. *Synthese* **18** (1968), 216-33
- Zbigniew Majewski: The place of physics in natural sciences. *Stud. Filozof.* **3** (1967), 203-20
- J. Margolis: Taylor on the reduction of teleological laws. *Inquiry* **11** (1968), 118-24
- G. Schlesinger: What does the denial of absolute space mean? *Aust. J. Phil.* **45** (1967), 44-60
- Howard Stein: On Einstein-Minkowski space-time. *J. Phil.* **65** (1968), 5-23
- C. Taylor: A reply to Margolis. *Inquiry* **11** (1968), 124-8
- John W. Thompson: Modes of thought in meteorology. *General Systems* **12** (1967), 201-7

Biological, Behavioural and Social Sciences

- E. M. Adams: Mind and the language of psychology. *Ratio* **9** (1967), 122-39
- Philip K. Bock: Three descriptive models of social structure. *Phil. Sci.* **34** (1967), 168-74
- M. Capek: Ernst Mach's biological theory of knowledge. *Synthese* **18** (1968), 171-91
- Robert R. Ehman: Causality and agency. *Ratio* **9** (1967), 140-54
- I. C. Jarvie: On theories of fieldwork and the scientific character of social anthropology. *Phil. Sci.* **34** (1967), 223-42
- Jaegwon Kim: Psychophysical laws and theories of mind. *Theoria* **33** (1967), 198-210
- Hugh Lehman: Are biological species real? *Phil. Sci.* **34** (1967), 157-67
- R. N. McLaughlin: Human action. *Aust. J. Phil.* **45** (1967), 141-58
- Torhy T. Segerstadt: Bemerkungen zur soziologischen Theoriebildung. *Köln. Z. Soziol. Sozial-Psychol.* **3** (1967), 429-39
- H. Szarski: The concept of progress in evolution. *Scientia* **103** (1968), 152-60
- J. R. Ullman: Artificial perception. *Scientia* **102** (1967), 476-87

AN INTRODUCTION TO THE PHILOSOPHY OF SCIENCE

D. W. THEOBALD

A stimulating account of the philosophical implications of scientific reason—and of the ways in which science and philosophy may fruitfully interact.

A University Paperback
Original 10s 6d
Hardback 30s

METHUEN

THE UNDECIDABLE

**Basic Papers on Undecidable Propositions,
Unsolvability Problems and Computable
Functions**

Edited by Prof. Martin Davis, New York University.

This self-contained source book is at once a standard reference work on the subject as well as a text or supplementary text in logic, philosophy and foundations of mathematics.

Kurt Gödel: On Formally Undecidable Propositions of the Principia Mathematica and Related Systems; On Undecidable Propositions of Formal Mathematical Systems (the famous 1934 Princeton lectures never before published, with corrections, emendations and a Postscript prepared for this anthology); Remarks Before the Princeton Bicentennial Conference on Problems in Mathematics; On the Length of Proofs; On Intuitionistic Arithmetic and Number Theory. Alonzo Church: An Unsolvability Problem of Elementary Number Theory; A Note on the Entscheidungsproblem. A. M. Turing: On Computable Numbers, with an Application to the Entscheidungsproblem; Systems of Logic Based on Ordinals. J. B. Rosser: An Informal Exposition of Proofs of Gödel's Theorem and Church's Theorem; Extensions of Some Theorems of Gödel and Church. S. C. Kleene: General Recursive Functions of Natural Numbers; Recursive Predicates and Quantifiers. E. L. Post: Finite Combinatory Processes, Formulation 1; Recursive Unsolvability of a Problem of Thue; Recursively Enumerable Sets of Positive Integers and Their Decision Problems; Absolutely Unsolvability Problems and Relatively Undecidable Propositions—Account of an Anticipation.

440 pages, with index, clothbound \$9.75

RAVEN PRESS, Box 52, Hewlett, N.Y. 11557

The British Journal for the Philosophy of Science

is one of the 45 learned journals published by Cambridge University Press.

A descriptive catalogue of Cambridge journals and a catalogue of books on the history and philosophy of science are available from the London and New York offices of the Press.

Cambridge journals may be ordered from a bookseller or direct from the publishers.

Cambridge University Press
Bentley House, 200 Euston Road
London, N.W.1
American Branch: 32 East 57th
Street
New York, N.Y. 10022

Pitman Books

Variational Principles in Dynamics and Quantum Theory

W. Yourgrau and S. Mandelstam

Third Edition 40s net

This celebrated monograph receives a new chapter on variational principles in hydrodynamics, which is, substantially, original work by the authors, and of great current interest in this field. Otherwise the book is unaltered and remains the best general account of variational principles in physics.

39 Parker Street London WC2

Philosophy and Phenomenological Research

AN INTERNATIONAL QUARTERLY

Edited by MARVIN FARBER

SEPTEMBER, 1968, VOLUME XXIX, No. 1

CONTENTS

I. ARTICLES

- Induction and Husserl's Theory of Eidetic Variation, by DAVID MICHAEL LEVIN
 Fiction and Phenomenology, by DONALD B. KUSPIT
 The Role of Imagination in Judgments of Fact, by D. C. S. OOSTHUIZEN
 A Critique of Contemporary American Philosophy, A Plea for Creativity, by ARTHUR W. MUNK
 The Discernment of Moral Attributes, by L. O. KATSOFF
 Deciding for God—The Bayesian Support of Pascal's Wager, by MERLE B. TURNER
 The Expression Theory of Avowals, by JAMES E. TOMBERLIN

II. DISCUSSIONS

III. REVIEWS, NOTES AND NEWS, AND RECENT PUBLICATIONS

Subscription prices \$6.00 per year for individuals; \$7.50 per year for institutions.

Separate issues, \$1.65 and \$2.00 respectively.

Postage and handling for back issues only.

A circular listing the main contents of the Journal since 1940 will be sent upon request

Address all communications to

PHILOSOPHY AND PHENOMENOLOGICAL RESEARCH

State University of New York at Buffalo, Buffalo, New York, 14214

MIND

A Quarterly Review of Psychology and Philosophy
 edited by Professor GILBERT RYLE

OCTOBER 1968

- | | |
|---|------------------------------|
| I. Legal Duties, Offences, and Sanctions | A. D. WOZZLEY |
| II. Aristotle's Conception of Syllogism | JAMES DUELLINGER |
| III. On the Connotation and Sense of Proper Names | CHRISTOPHER KIRWAN |
| IV. A Principle about 'About' | D. HOLDCROFT |
| V. A Definition of Memory | E. M. ZEMACH |
| VI. Propositions and Eternal Sentences | CHARLES SAYWARD |
| VII. Discussions | |
| Must we talk about "Is" and "Ought"? | DOROTHY MITCHELL |
| Categories of Psychological Explanation | D. R. FINN |
| How one could tell were a bee to guide his behaviour by a rule | |
| Hart's Primary and Secondary Rules | ARTHUR W. COLLINS |
| Professor Goddard and the Simple Theory of Types | K.-K. LEE |
| Are Rules of Moral Thinking Neutral? A Note on Liberty and Equality | M. D. RESNIK |
| A Note on "Time, Truth, and Modalities" | DAVID SIDORSKY |
| Choosing as Doing | JAMES CARGILE, GEORGE THOMAS |
| Religious Morality: A Reply to Flew and Campbell | ROGER HANCOCK |
| Raziel Abelson on "Because I Want To" | PATTERSON BROWN |
| On the Possibility of Rational "Inconsistent" | JON PASHMAN |
| Beliefs | |
| Malcolm on Mind and the Human Form | E. S. EDGINTON |
| VIII. Critical Notices | L. D. HOULGATE |
| Stephen Körner, <i>Experience and Theory: An Essay in the Philosophy of Science</i> | |
| ROGER BUCK | |
| Brian Barry, <i>Political Argument</i> : J. F. DAY | |
| IX. New Books | |
| X. Notes | |

Annual Subscription 25s, post free; single copy 7s 6d, plus postage

Published for MIND ASSOCIATION by

BASIL BLACKWELL

49 BROAD STREET OXFORD

INQUIRY

An Interdisciplinary Journal of
Philosophy and the Social Sciences
Edited by ARNE NAESS

Volume 11

Number 3

Autumn 1968

God, Guilt, and Logic: the Psychological Basis of the
Ontological Argument

Lewis S. Fener

Principles of Polemic in Russell

Harry Ruja

The Propositional Logic of Ordinary Discourse

William S. Cooper

Comte and the Idea of Progress

Leslie Sklair

Discussion:

Wittgenstein and Self-Description

V. M. Hope

Review Discussion:

Professor Narveson's Utilitarianism

T. L. S. Spriggs

*Published quarterly by Universitetsforlaget, P.O. Box 307,
Blindern, Oslo 3, Norway*

Annual Subscription: 35s., \$6.00, postage included

NOW AVAILABLE

Twenty year
Cumulative Index
1947-1967

THE REVIEW OF METAPHYSICS

The Twenty year cumulative index covering the first twenty volumes of *The Review of Metaphysics* is divided into three sections: a subject index; an author index; and an index of books that have been examined in *Critical Studies*. \$3.00

Orders can now be placed for the Index.

For further details write to

THE REVIEW OF METAPHYSICS,
214 Lyman Beecher Hall, Haverford College,
Haverford, Pa. 19041.

JOURNAL OF THE HISTORY OF IDEAS

A Quarterly devoted to Intellectual History

VOLUME XXIX

OCTOBER-DECEMBER 1968

NUMBER 4

TABLE OF CONTENTS

ARTICLES

- Homeric Psychology and the Oral Epic Tradition *Joseph A. Russo*
Paradoxes of Solomon: Learning in the English Renaissance *M. Haltway*
The Faerie Queene and Cosmogonies of the Near East *B. G. Feinstein*
"Interest Will Not Lie": A 17th-Century Political Maxim *J. A. W. Gunn*
Galileo, Ficino, and More's *Psychathanasia* *C. A. Staudenbaur*
Perry Miller's *New England Mind* *Gene Wise*

NOTES

- A Note on Stating the Socratic Paradox *John J. Mulhern*
Buber's *Dialogue* and Gandhi's *Satyagraha* *V. V. Ramana Murti*
Justinian I and the *Descensus* *Heleen Forrest*

REVIEW

- Encyclopedia of Philosophy*, edited by Paul Edwards Reviewed by *Phillip P. Wiener*

Subscription Rates (outside U.S.A. post paid): \$7.50 for one year;
\$2.00 for single issue

Payable to

JOURNAL OF THE HISTORY OF IDEAS

The City College, New York, N.Y. 10031

FOUNDATIONS OF LANGUAGE

International Journal of Language and Philosophy

Editorial Board: Morris Halle (MIT), Peter Hartmann (Münster/W), K. Kunjnni Raja (Madras), Benson Mates (Univ. of California), J. F. Staal (Amsterdam), Pieter A. Verburg (Groningen), John W. M. Verhaar (Nijmegen)

VOLUME IV, No. 2, MAY 1968

Articles: Wallace L. Chafe: Idiomaticity as an Anomaly in the Chomskyan Paradigm—Joseph Margolis: Quine on Observationality and Translation—Thomas E. Patton: Syntactic Deviance—Franz W. Dörge: Publizistikwissenschaft und Sprache als ihr Formal-object—*Discussion:* Meaning, Regular and Irregular (J. F. Staal)—*Reviews by:* William Kneale, Richard M. Gale, Anatol Rapoport, W. Wexsel, Jacek Fisiak, Daam M. van Willigen and Ronald W. Langacker—*Short Notices—Publications Received for Review*

FOUNDATIONS OF LANGUAGE is published quarterly

Annual subscription price f42.—(\$11.75)

D. REIDEL PUBLISHING COMPANY/DORDRECHT - HOLLAND

dialectica

Revue trimestrielle de philosophie de la connaissance

- s'attache à promouvoir les principes directeurs de la *méthodologie ouverte*. Image vivante du développement de la philosophie ouverte, elle publie sur cette dernière et ses thèses essentielles des éditoriaux, des articles de fond, les comptes rendus des débats et des discussions sur ses tendances et ses engagements actuels
 - organise des débats et publie des enquêtes sur les questions essentielles de la connaissance moderne (voir par exemple les numéros sur l'idée de complémentarité avec la participation de cinq prix Nobel de physique, sur les fondements de la psychologie, du calcul des probabilités, sur la notion d'analogie, le fait et le droit, l'épistémologie de la physique quantique)
 - se prête à la remise en cause des bases des disciplines abstraites et le fondement des disciplines expérimentales
 - commente les publications récentes et accorde l'hospitalité de ses colonnes aux discussions de principe à leur sujet
 - appelle à la collaboration tous ceux qui entendent participer à la synthèse et à l'intégration de l'ensemble des connaissances modernes
 - procède actuellement à une enquête sur la question brûlante de la rénovation de l'enseignement des mathématiques, de la physique et de la biologie
- dialectica en un mot est l'organe vivant et la libre tribune de la *méthodologie* et de la *philosophie ouvertes*

Comité directeur: FERDINAND GONSETH, Lausanne
PAUL BERNAYS, Zurich
H. KÖNIG, Wabern/Berne
P. B. FILET, Lausanne

Rédaction: 12, chemin du Muveran, 1000 Lausanne (Suisse) Vente et abonnements: W. Roesch & Cie, Monbijoustrasse 9, 3000 Berne (Suisse)

Prix: Suisse Fr 9.-le numéro Etranger: Fr 10.-le numéro
Abonnement: Suisse Fr 30.- Etranger: Fr 35.-

Revue Internationale de Philosophie

Editorial Board: Profs. Ph. DEVAUX, H. GOUHIER, Ch. PERELMAN and Mrs. Simone SERVAIS

Secretary: 99 avenue de l'Université, Bruxelles 5

Each number is devoted to a particular movement, a particular philosopher, or a particular problem. As complete a bibliography as possible is given.

Truly international in spirit, the *Review* chooses its collaborators among scholars of all countries. The articles are written in English, French, German, Italian, or Spanish.

Publishes 4 numbers annually. (Europe: 350 Belgian francs; other countries: 400 Belgian francs.)

Coming numbers for 1968 and 1969

Philosophie du Langage au XVIII^e: Siècle, Eug. Dupréel, L'Humanisme, L'Analogie, Wittgenstein

" SCIENTIA "

INTERNATIONAL REVIEW OF SCIENTIFIC SYNTHESIS

(1968 - 62nd Year)

Scientific Committee

G. ABETTI - G. B. BONINO - P. CALDIROLA - L. CALIFANO - R. CANESTRARI -
G. COLONNETTI - B. DE FINETTI - G. MERLA - G. MONTALENTI - E. PADOA - G. SANSONE
A. SESTINI - V. SOMENZI - S. TONZIG.

Editor N. BONETTI

" SCIENTIA " publishes original articles by leading workers in all the fields of science and philosophy, from all countries.

It is a magazine of synthesis, that is, of the convergence of the various branches of knowledge.

It is a magazine of multidisciplinary information; it promotes the " bridge " disciplines, which are the most active in the progress of research.

It encourages the building of a culture fully adherent to modern scientific and philosophic standards.

Published every two months, each issue containing approximately 200 pages. Articles are published in the original language of the author i.e. English, Italian, French, Spanish and German. Every issue contains a supplement with a French translation of the non-French articles. Each issue contains: Book reviews - Analysis - Chronicle.

SUBSCRIPTION RATES: Italy Lire 15.000 - Abroad \$32.00 or £13
Current issue: Italy Lire 3.000 - Abroad \$6.50 or £2.

Back issues: Italy Lire 500 - Abroad \$1.50 or 11s. (in stamps).

"Scientia" representative in U.K.: Atlas Publishing & Distributing Co. Ltd. -
334 Brixton Road - London S.W.9

" SCIENTIA " - Via Donizetti, 55 - 20122 Milano (Italy)

THE BRITISH JOURNAL FOR THE HISTORY OF SCIENCE

CONTENTS OF VOL. IV, PART I, No. 13

JUNE 1968

Main Articles:

The Background to the Discovery of Dulong and Petit's law
The cause of refraction in Medieval Optics
The third law in Newton's mechanics

ROBERT FOX
DAVID C. LINDBERG
RODERICK W. HOME

Notes and Communications:

Lavoisier's early career in science. An examination of some new evidence.

Babbage and Moll on the state of science in Great Britain. A Note on a Document

Chemistry in space and the complex atom

J. B. GOUGH
NATHAN REINGOLD
W. V. FARRAR

Book Reviews:

**PUBLISHED BY THE BRITISH SOCIETY FOR THE
HISTORY OF SCIENCE**

Registered Office;

SCIENCE MUSEUM, SOUTH KENSINGTON, LONDON, S.W.7

HISTORY OF SCIENCE

An Annual Review of Literature, Research and Teaching

Edited by A. C. CROMBIE and M. A. HOSKIN

CONTENTS OF VOLUME 6, 1967 42s.

ARTICLES:

- Rosalind C. H. Tanner*—The Study of Thomas Harriot's Manuscripts: I. Harriot's Will
Jon V. Pepper—The Study of Thomas Harriot's Manuscripts: II. Harriot's Unpublished Papers
A. P. Youskevitch—Recherches sur l'histoire des mathématiques au Moyen Age dans les pays d'Orient
D. T. Whiteside—A Face-lift for Newton: Current Facsimile Reprints
J. A. Lohne—The Increasing Corruption of Newton's Diagrams
John R. Levene—Sources of Confusion in Descartes's Illustrations

ESSAY REVIEWS:

- Carl B. Boyer*—The Making of a Mathematician
C. Webster—The Origins of the Royal Society
A. G. Debus—Alchemy and the Historian of Science
J. S. Wilkie—Preformation and Epigenesis
Richard S. Westfall—The Science of Optics in the Seventeenth Century
W. H. Brock—The First Principles of the History of Chemistry
Erwin H. Ackerknecht—The Total Harvey

Published by

W. HEFFER & SONS LTD., Cambridge, England

SYNTHESE

*An International Journal for Epistemology, Methodology and
Philosophy of Science*

Editor: JAAKKO HINTIKKA

University of Helsinki and Stanford University

Board of Consulting Editors: David M. Armstrong, Yehoshua Bar-Hillel, Arthur W. Burks, Robert S. Cohen, Donald Davidson, Dagfinn Føllesdal, Hans Freudenthal, Béla Juhos, Jerrold J. Katz, Imre Lakatos, G. Nuchelmans, Wesley C. Salmon, Wolfgang Stegmüller, Erik Stenius, Patrick Suppes, Klemens Szanlawski, H. Törnebohm, Marx Wartofsky.

Contents of Volume 18, No. 1, January 1968

- The Structure of Logical Probabilities by *Jens Erik Fenstad*
The Whole Truth by *Richard C. Jeffrey*
On the Confirmation of Hypotheses about Regions of Existence by *Håkan Törnebohm*
Causal Connection by *William Ruddick*
A Generalization of the Principle of Causality, Which Makes It Applicable to Evolutionary Systems by *Claude Hillinger*
Causality Requirements and the Theory of Relativity by *Peter Haver*
No more Discovery in Physics? Review Essay by *Joseph Agassi*

Reviews

Books Received

Annual Subscription Price Dfl. 60.—(\$17.00)

Personal Subscription Price Dfl. 35.—(\$9.80)

D. REIDEL PUBLISHING COMPANY
DORDRECHT-HOLLAND

VOLUME X

JUNE 1968

No. 1

RATIO

edited in co-operation with

H. B. Acton (London), P. Bernays (Zürich), A. Church (Princeton),
H. Cherniss (Princeton), G. Henry-Hermann (Bremen), W. C.
Kneale (Oxford), E. Nagel (Columbia), K. R. Popper (London)
and J. W. N. Watkins (London)

by

Stephan Körner (Bristol)

W. STEGMÜLLER

Towards a Rational Reconstruction of Kant's
Metaphysics of Experience, Part II

PETER WINCH

Wittgenstein's Treatment of the Will

JÜRGEN MITTELSTRASS

Predication and Recurrence of the Same

WOLFGANG YOURGRAU AND
CHARLES WORKS

A New, Formalized Version of the Verifiability
Principle

RICHARD S. WESTFALL

Huygens's Rings and Newton's Rings: Periodicity
and Seventeenth Century Optics

REVIEW:

GRAHAM BOND

Lewis White Beck: Studies in the Philosophy of
Kant

Price 15s. net

Annual Subscription: 27s. 6d. net

BASIL BLACKWELL · OXFORD

THE AUSTRALASIAN JOURNAL OF PHILOSOPHY

EDITED BY GRAHAM NERLICH

VOL. 46

AUGUST, 1968

No. 2

ARTICLES:

G. FREGE

On the purpose of the Begriffsschrift

IAN HACKING

A theory of indefinite descriptions

HUGH S. CHANDLER

Persons and predicability

JAMES W. CORNMAN

Private languages and private entities

NICHOLAS MAXWELL

Understanding sensations

JORDAN HOWARD SOBEL

Rule Utilitarianism

CRITICAL NOTICES

K. W. RANKIN

W. D. Joske: *Material Objects*

KEITH CAMPBELL

C. F. Presley (Ed.): *The Identity Theory of Mind*

REVIEWS

BOOKS RECEIVED

Published in May, August and December; Annual Subscription 41s 10d Sterling (32s 6d to members
of the Australian Association of Philosophy, which you are invited to join). A sample copy and
other information may be obtained from the Department of Philosophy, University of Sydney, N.S.W.

THE BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

Committee 1968-69

President: Mr J. L. Mackie

Vice-President: Dr G. J. Whitrow

Hon. Secretary: Mr R. Harré

Hon. Treasurer: Dr W. Mays

Hon. Secretary of the Northern Group: Dr N. Wetherick

Ordinary Members: Dr J. Giedymin, Mr A. E. Musgrave, Mr H. D. Sluga

EDITORIAL BOARD OF THE JOURNAL

Professor J. D. Bernal

Professor R. B. Braithwaite

Professor H. Dingle

Professor Sir John Eccles

Professor Sir Harold Jeffreys

Professor W. C. Kneale

Professor S. Körner

Professor M. Polanyi

Professor Sir Karl Popper

Earl Russell

Dr G. J. Whitrow

Mr L. L. Whyte

Professor J. H. Woodger

The British Journal for the Philosophy of Science is the official organ of the Society, whose purpose is described by the following statement from its Constitution and Rules :

‘ The purpose of the Society is to study the logic, the method and the philosophy of science as well as those of the various special sciences, including the social sciences.’

The Society holds meetings, at which papers are read and discussed, once a month during the academic year. These are held, by kind permission, in University College, London.

Ordinary membership is open to all persons who have been proposed by two members, approved by the Committee, and elected at an ordinary meeting. The annual subscription is 20s. Those who wish to join should write to the Hon. Secretary at Linacre College, Oxford.

The Northern Group of the Society meets regularly at various centres in the north of England. Those who wish to join it should write to the Hon. Secretary of the Northern Group at the Undergraduate School of Studies in the Social Sciences, University of Bradford, Bradford 7.

The British Journal for the Philosophy of Science

ARTICLES

- 185-210 JEFFREY BUB Hidden variables and the Copenhagen interpretation—a reconciliation
211-223 LARRY WRIGHT The case against teleological reductionism
225-246 CARL-ERIK SARNDAL Some aspects of Carnap's theory of inductive inference

DISCUSSIONS

- 247-251 WILLIAM L. HARPER AND HENRY E. KYBURG The Jones Case
251-253 PAUL K. FEYERABEND A note on two 'problems' of induction
254-256 ERROL E. HARRIS Simultaneity and the future
257 J. E. J. ALTHAM A note on Goodman's paradox

258 ANNOUNCEMENT

259-274 REVIEWS

275-280 RECENT PUBLICATIONS

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, N.W.1.

American Branch: 32 East 57th Street, New York, N.Y. 10022

Single issues 20s. net in U.K. (\$2.75 in U.S.A.)

Subscription price 60s. net in U.K. (\$9.50 in U.S.A.)

Volume 19 Number 4

February 1969

The British Journal for the Philosophy of Science



CAMBRIDGE UNIVERSITY PRESS

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

Editors: MARY B. HESSE, D. H. MELLOR

Review Editors: B. M. FLOOD, G. M. K. HUNT

Subscriptions

The *Journal* is published quarterly. The subscription price is 60s. net in U.K. (U.S.A. \$9.50) per volume post free; single issues 20s. net in U.K. (U.S.A. \$2.75) plus postage. The subscription is 40s. (U.S.A. \$7.00) for members of the British Society for the Philosophy of Science, who should order direct from the publisher. Orders may be sent to any bookseller or direct to Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1. Orders from U.S.A. and Canada should be sent to Cambridge University Press American Branch, 32 East 57th Street, New York, N.Y. 10022, U.S.A. Second class postage paid at New York, N.Y.

Notes to Contributors

Articles and correspondence should be sent to the Editors, and books for review to the Review Editors, at the Whipple Museum, Free School Lane, Cambridge CB2 3RH. Articles should not normally exceed 5,000 words, and should be accompanied by abstracts of about 100 words which will then appear in *Dialectica*, *Synthese* and *Philosophy of Science*.

All contributions should be typed in duplicate on one side only with wide margins and double-spaced. Footnotes, which should be as few as possible, should be typed separately, double-spaced at the end of the text. They should be referred to by consecutive numbering throughout the text, in which they will be inserted in page proof. Bibliographical references should be listed alphabetically at the end, and reference made to this list in the text rather than in footnotes. The style of references should in future follow the Harvard system used in this issue.

Long quotations are printed in small type without quotes, and should be so marked in the typescript. Diagrams, tables, and illustrations should be on separate sheets, with their desired position in the text clearly indicated. Symbolic formulae appearing as separate lines should be clearly distinguished in the text, and foreign letters (Greek, German, etc.) indicated in the margin. On matters of style and layout, authors should consult CHAUNDY, T. W. *et al.* (1964). *The Printing of Mathematics*. London: Oxford University Press.

Offprints of articles are ordered when corrected galley proofs are returned. 50 are supplied free; the price of extra copies is given on the order form.

Advertising

Particulars of advertising in the *Journal* may be obtained from Cambridge University Press, Bentley House, 200 Euston Road, London, N.W.1.; or in U.S.A. and Canada, Cambridge University Press American Branch, 32 East 57th Street, New York, N.Y. 10022, U.S.A.

Vagueness, Inexactness, and Imprecision

by R. G. SWINBURNE



I

There is often uncertainty about whether some predicate applies to some physical object (or state), whether the object is green or blue, exactly 2 cm long or less than 2 cm in length. The uncertainty may be revealed either by observers hesitating over application or by different observers making conflicting judgments. The existence of this uncertainty raises two interesting questions which have received some discussion in recent writing. The first is the question why the uncertainty exists, and the second is the question whether such uncertainty is a necessary feature of any language in any world. Black (1949) and Waismann (1945), and, more recently, Körner (1966), Mellor (1965, 1966 and 1967), and Kumar (1967) have all discussed these issues but their results overlap and conflict. I wish in this paper to give a comprehensive answer to the questions, bringing the results of these authors into perspective, disagreeing with some of their points and elucidating others.

Let us begin by considering briefly the general question of what are the criteria for correct application of terms. There are two ways in which terms derive their meaning and from which are derived standards for their correct use. One is by their analytic relations to other terms and the other is by the phenomena which are standard instances and standard non-instances for their correct application. In other words, there are syntactic and semantic criteria for the correct use of terms.

Many technical scientific terms, such as 'cistron', 'half-life', and 'magnetic susceptibility', have only syntactic criteria for their correct application. The magnetic susceptibility of a substance is—by definition—the intensity of magnetisation induced in a unit volume of the substance by a unit magnetic field. If, however, a term is an empirical term, semantic criteria must be provided for the terms in its definition, or for terms occurring in definitions of these terms, or yet more remotely in definitions of the latter. Ultimately, if it is to be used for describing the world, semantic criteria must be provided for the use of a term. By the semantic criteria

Received 15 July 1968

I am most grateful to Hugh Mellor and to Professor Alan White for their helpful criticisms of earlier versions of this paper.

an object will be *f* if it is similar in some respect to standard examples of objects which are *f* and dissimilar in that respect from the standard examples of objects which are $\sim f$. For many terms criteria of correct application are given both by rules connecting their use with the uses of other terms (e.g. 'a chair is an object which is normally used for sitting on by one person and has a back rest') and by standard examples (viz. pointing out objects which are chairs and distinguishing them from similar objects which are not). But in the end the other terms themselves have to be cashed by examples.

2

Now, given this general answer to the question of what are the criteria for correct applications of terms, let us consider why there is uncertainty about the applicability of terms. One obvious source of such uncertainty is that different users of language have different defining criteria for correct application of some term. I will call this source of uncertainty terminological vagueness. Such vagueness may arise either because men have different syntactical criteria for the use of a term or because they regard different cases as standard cases of correct or incorrect application of a term. An example of the first source of terminological vagueness is the following example, adapted from one of Mellor's. One man may mean by the population of a town in some year the average number of permanent residents of houses within its boundaries on any one day during that year; and another man may mean the average number of people who work or reside within its boundaries on any day during the year. Disagreement between them about what is the population of the town in some year would then be due to the vagueness of the term.

An example of vagueness arising for the second reason is given by Kumar (1967, p. 219). He makes the point that 'a person acquainted with "orange" besides "red" and "yellow" will apply "non-red" to certain objects to which one not acquainted with "orange" will apply "red"'. The two people will have a different standard example of red and non-red objects and because of this will designate other objects as red or not red in different ways. That some disagreement about correct application of a term arises from different users of language having different criteria for its use can be brought out by studying their applications of the term on other occasions and asking them about their criteria for its use. When a term is used in two or three distinct ways, we say that it is ambiguous, but if it is used in a whole spectrum of ways we often simply say that it is vague. In either case I shall call it vague. It is natural to say that when a term is vague in this

sense, then the term denotes more than one concept; each user of language with distinct criteria for the use of a term uses it to denote a distinct concept.

Even if terms are vague, it is not necessary that there ever be in fact any uncertainty about their applicability to any particular object, for the different criteria for the correct application of terms may give the same result for any actual object considered by observers. But if a term is vague, it is logically possible that the different criteria for its applicability might give rise to uncertainty about whether or not the term applied to some object.

When referring to the criteria for the application of a term I was of course referring to the defining criteria or primary tests for its correct application. A primary test for an object being f is a member of the smallest set of tests, semantic or syntactic, such that, whatever the result of other tests, if an object is f by all these tests, then it is clearly f , and if it is $\sim f$ by all these tests, then it is clearly $\sim f$. The primary tests give the meaning of the term. Secondary tests are tests used because it is believed that their results are correlated with the results which would be given by primary tests, and they are used because of the difficulty of applying primary tests.

Now even when users of language have the same primary tests for a term, and so any term is used to denote one and only one concept, uncertainty about the application of terms may still remain, and there are exactly two possible sources of this. Either the concept is too blurred so that it is unclear whether or not it fits an object; or while the concept is clear enough, observers are in no position to know for certain whether or not it fits—they may not be good enough observers or have adequate enough instruments. Following Mellor, I will term the former source of uncertainty about correct application of terms 'conceptual inexactness', and I will term the latter source 'practical difficulty'. A concept will be inexact if it is logically possible for an object to be such that the primary tests, however rigorously executed, do not determine definitely whether or not it falls under the concept. The inexactness of the concept will be responsible for uncertainty about whether or not it applies to some object if primary tests have failed to determine definitely whether it applies, and would fail to do so even if applied more rigorously.

Now, bearing these points in mind, consider how investigators ascertain if a concept applies to an object. We will then see how they may become uncertain about a concept's application and in different cases practical difficulty or conceptual inexactness may be responsible for their uncertainty. First, investigators may use secondary tests. If these give conflicting or unclear results, the investigators will be uncertain whether or not the concept applies to the object. This will be the result of practical difficulty. The

trouble is not that the concept is inexact but that we have used second-best tests for its application. We might have two different secondary chemical tests for whether or not a substance was highly inflammable, the primary test of course being to light it and see if it burns rapidly. If these secondary tests give *conflicting* results, at least one of them is misleading. The solution is to use another test better correlated with the primary test, or to use the primary test itself. A secondary test may also give an *unclear* result; for example, if it does not allow us to estimate whether the value of some quantity is greater than or less than a specified amount. A dot on the radar screen for locating ships may hover on the 15-mile line and so not allow us to say if the concept 'further away than 15 miles' applies to the ship represented by the dot. The solution is to use a more sensitive instrument, or to measure the distance with a tape (viz. to use the primary test).

Mellor (1965, p. 109) claims that 'any law into which any . . . [empirical] . . . term enters can be used to provide at least part of an operational definition of it', viz. any secondary test for the applicability of a term can become a primary test. This is true, but at any instant one can make a distinction between the primary and secondary tests for the application of a term. If the primary tests for the application of a term change, then the term has changed its meaning, denotes a different concept. Further, there is no necessity for any secondary test to become a primary test. A second point relevant here in criticism of Mellor is this. Mellor seems, in the various articles cited, to suggest, though he does not explicitly state, that if tests of different kinds for the application of a concept give different results, then the concept is inexact. My analysis suggests that this does not follow unless the tests are primary tests.

Suppose, however, that investigators use only primary tests for whether or not a concept f applies to some object. They will be uncertain about whether the concept applies to the object if the tests give different or unclear results. There may be many primary tests, syntactic and semantic, for the applicability of a concept and there may be various sets of these such that if all members of a set are satisfied, then the object certainly is f and if no members are satisfied, then the object certainly is $\sim f$, but if some of the tests are satisfied and others not satisfied, it is not certain whether or not the object is f . Thus there are many different primary tests for whether an object is some chemical substance. A gas is chlorine if:

- (i) 1 litre of the gas at standard temperature and pressure weighs *c.* 3.1667 gm.
- (ii) its critical temperature is 140°.

- (iii) its boiling point is -3.36° .
 - (iv) the melting point of its liquid form is -102° .
 - (v) when it is mixed with water and some hydrochlorous acid is formed, some hydrochloric acid is also formed, and . . .
- various other standard chemical tests are satisfied.

Clearly if the gas satisfies all these tests it is chlorine, and if it satisfies none of them it is not. It is not necessary that it satisfy all the tests for it to be considered chlorine—if it satisfied a considerable majority of them but failed to satisfy a few, it would still be considered chlorine. If it failed to satisfy a considerable majority but satisfied a few, it would not be considered chlorine. If it satisfies or does not satisfy one or two tests, but the others have not been applied, this is often reasonable grounds for supposing that it will satisfy, or not satisfy as the case may be, the others. Yet the satisfaction of some half of the tests including chemical and physical ones and non-satisfaction of the other half would leave it unclear whether or not the gas was chlorine. If the tests applied are all the primary tests that there are, then no further tests can settle the matter. Hence the inexactness of the concept is responsible for the uncertainty about whether it applies. Yet if only a small minority of the tests have been applied, the uncertainty could be resolved by a decisive majority of the rest. In that case we do not know whether the uncertainty is due to conceptual inexactness or practical difficulty. There are, however, other tests such as smelling and tasting the gas to determine whether it is chlorine, the results of which would not be considered relevant if we have the results of primary tests. They are therefore secondary tests. For other concepts there are few primary tests and many secondary tests in common use. The one primary test for distance is given by measurements along a straight line with a tape or ruler. There are many different tests for measuring the distance of distant stars—by parallax, by cepheid variables, etc. But if we could measure the distance with a ruler, we would know what the distance was, whatever the results of the latter tests. They are therefore secondary tests.

So then if primary tests are divided fairly evenly for or against an object being an instance of some concept it will be uncertain whether or not it is. If results of other primary tests could settle the issue, then practical difficulty may be responsible for the uncertainty. Otherwise the inexactness of the concept is responsible. Mellor (especially 1966) has drawn attention to this kind of inexactness. He considers the concept of a tributary, such that a river branch is a tributary if it is shorter in length than the branch which meets it and also if it has a smaller volume flow than the latter branch. If however, the two criteria give conflicting results, such that by one

a branch is a tributary and by another it is not, it is uncertain what we ought to say. Alston (1964, p. 88) considers the concept of a religion and gives nine criteria for a cultural entity being a religion, such as that its members believe in supernatural beings (gods), perform ritual acts, pray to the gods, and so on. If the entity has all these characteristics, then it is a religion; and if it has none of them it is not; but it is uncertain what we ought to say if it has only half of them, and not the other half. In Körner's terminology (1966, p. 28) such a concept would, I suspect, be an 'incomplete' one, and an object which satisfied half of the criteria but not the other half would be an 'undecidable' candidate for being an instance of the concept. Again, as with the similar case of vague terms, the existence of distinct criteria makes it logically possible to be uncertain if a term applies to an object, although in fact the criteria may give the same results for all actual objects.

Uncertainty may however arise, not because different primary tests give different results, but because one primary test gives unclear results. If there is more than one primary test for the application of a concept, and some give unclear results, but others give the same clear result, we would normally take the verdict of the latter. Trouble arises, however, when all primary tests, perhaps the one and only primary test, give unclear results.

This unclarity may have any of several sources. One is that the primary test may not have been fully or rigorously executed, and for that reason the investigator is uncertain—within a range—about the result of the test. Thus I may try to ascertain the colour of an object by comparing it with the standard patches in bad light. My experience of how objects in bad light differ in apparent colour from those in good light allows me to affirm with confidence that the object is not red or pink or white, but it might be either blue or green. This uncertainty is evidently the result of practical difficulty. It would disappear with an improvement in the light.

However, in this connection there is a crucial difference between concepts of two kinds: those, the primary tests for the application of which can logically be executed with full rigour, and those, the primary tests for the application of which can logically only be executed with ever increasing but never perfect rigour. The former I will term 'A-concepts' and the latter 'B-concepts'. The primary tests for A-concepts are made with the naked eye or with instruments of only a certain degree of elaboration. Most concepts of ordinary language are A-concepts—'green', 'chair', 'hill', 'river', 'opaque', 'scented' are all terms denoting A-concepts. We do not need instruments to tell whether or not an object is a chair or a hill. Many scientific concepts however are B-concepts, as for example, are the

concepts 'heavier than 5 gm' and 'exactly 9 volts'. We get more and more reliable results about whether an object weighs more than 5 gm, or a potential difference is exactly 9 volts by using more and more elaborate scales and voltmeters, and there is no limit to this process. Household scales and laboratory scales are constructed on the same principles and therefore using laboratory scales rather than household scales does not constitute applying a different test in the sense in which tests earlier distinguished are different from each other. Yet the laboratory scales have been more carefully and finely calibrated than the household scales and their results are therefore more reliable. Yet there is no limit to the care one can take in calibrating, in eliminating interfering factors etc., and so no limit to the reliability of scales. The effect of ever more rigorous tests will be continually to narrow the range of objects, actual or possible, about which it is uncertain whether or not the concept applies to them.

Now any uncertainty about the application of an A-concept arising from inadequate use of primary tests can be eliminated by using the tests more rigorously. This uncertainty is therefore due to practical difficulty, as in the example of ascertaining colour in a bad light. For B-concepts the position is different. The effect of more and more rigorous tests is continually to narrow the range of objects, actual or possible, about which it is uncertain whether or not the concept applies to them. Yet it is a consequence of this that, with a test of any given rigour, it is logically possible for objects to be within the range of uncertainty, and for this reason any B-concept is inexact. However for many B-concepts there is no need for the logical possibility to be realized. The narrowing of the range of uncertainty may eventually leave no actual objects inside it. Thus while an ordinary pair of scales might not allow us to say with certainty of all objects whether or not the concept 'heavier than 5 gm' applies to them, a pair of laboratory scales might settle the matter. But uncertainty about the application of all B-concepts could not be removed in this way, for among B-concepts are the concepts of exact values of such measurable properties as mass and length. They differ from other B-concepts in that they measure values of properties on a dense scale. A non-dense scale is a scale such as the scale of integral numbers, on which for every value there is a next value. A dense scale is a scale, such as the scale of rational numbers, by which measurements are normally made, on which for any value there is no next value. Now increasingly rigorous measurements narrow the range of uncertainty about the application of any given B-concept, and thus, for any given object, they narrow the range of B-concepts about which it is uncertain whether or not they apply to the object. Now only one exact value of some particular property—e.g. of length or mass—can apply to

any given object. If an object is 3 cm long, then it cannot be of any other length as well. Hence the increasingly elaborate measurements can only make it certain that a particular value of length or mass pertains to the object, if they make it certain that no other values pertain. The effect of narrowing the range of exact values, about which it is uncertain whether or not they belong to the object, will simply be, while more than one value remains, to show of more and more values that they do not belong to the object. The range within which it is uncertain which is the exact value of some concept of measurement is known as the 'imprecision' of the measurement. Now if measurements are made on a non-dense scale (as are, for example, measurements of the number of soldiers in a field or molecules in a volume of a substance), increasingly elaborate instruments may so narrow the imprecision as to leave only one value within the possible range. But if a dense scale is used, this can never happen, for on a dense scale there are within as small a distance as you like of any value, an infinite number of values. So only if increasingly elaborate tests made possible discrimination to an infinite degree of accuracy could they leave only one value as the true value of the property. Yet the supposition that men could discriminate to an infinite degree of accuracy is nonsensical, for in that case men would have to agree in their judgments when a body was, say, exactly 100°C in temperature and when it had any of the infinitely many values within any, however small, finite interval of 100°C . Yet such judgments could not be made, because even to utter, let alone agree on the application of, some of the infinite number of the values within that range, would take an infinite time. Therefore there is a necessary imprecision in the ascription of an exact value of a measurable concept whose possible values form a dense set. It is a practical matter how big the interval of imprecision is; it is a logical matter arising from the nature of the concept involved that there is an interval.

The second source of unclarity in the result of a primary test (and so uncertainty about the applicability of the relevant concept to an object) is the disagreement of observers about what the apparent result is. Thus suppose the terminological vagueness of colour terms to be removed, and colour terms defined by standard patches on a colour chart. An object will then be green if it resembles in colour the standard patch of green more than it resembles any of the other standard patches, and so on. Now even if the colour terms are defined in this rigorous kind of way, men will often be uncertain about the applicability of colour concepts, for often some observers will say that an object appears to resemble in colour one patch more than any other patch, and other observers say that it appears to resemble a different patch more than any other patch. Now the colour

of an object is that colour which it appears to have to a considerable majority of observers. If a small minority of all observers disagree with the vast majority about what is the apparent colour, then we may say that they do not discriminate colour correctly. But if a considerable minority disagree with the majority, then the colour of the patch is indeterminate.

Now if a number of observers disagree in fairly even proportions about the apparent colour of an object, then only if they are not typical of all observers, is the dispute about what is the real colour soluble. If the dispute arises because untypical observers are making the observations, the consequent uncertainty about whether the predicate belongs to the object is due to practical difficulty. If the observers are typical, the uncertainty is due to the inexactness of the concept. The concept is such that we cannot apply it in a clear way to all cases.

Thirdly, the unclarity may arise because of the different results of different applications of the test (say using different instruments, or on different occasions—when we have good reason to believe that the object does not change in the meantime in respect of the property in question). Thus, we might measure the wave-length of light from a distant star at intervals of a few seconds, and the readings might differ. Or we might use several spectrometers to measure it, and find that their readings differed. Now if this difference between the results can be removed by using more elaborate (larger, more carefully calibrated, etc.) instruments, then clearly the uncertainty was the result of practical difficulty. If no improvement in instruments would make for agreement between the results of tests then the concept is inexact.

Fourthly, the final result, agreed by all observers, of all applications of the primary tests may simply be that the object is a border-line case for being *f*. We can easily see how this situation can arise for A-concepts by modifying the colour example described earlier. As before, let colour be defined by standard patches, and the colour of an object be the colour of that patch to which it appears to most observers to be most similar. Now the colour of some object may appear to all observers to be midway between that of two patches, say 'blue' and 'turquoise'. Then the colour of the object will be border-line between 'blue' and 'turquoise'. It is the inexactness of the concept which is at fault in this kind of case—the concepts 'blue' and 'turquoise' are too blurred to fit exactly the colour of the object in question. Clearly for any set of colour terms, defined by standard patches, it makes sense to suppose that there be an object mid-way between the colour of any two standard patches, and hence the existence of border-line cases for the applicability of any colour terms is always logically possible. Colour concepts are therefore necessarily inexact. It is

on this kind of conceptual inexactness that Körner has concentrated in his recent book (1966), and he understands by conceptual 'inexactness' merely inexactness arising in this way, when primary tests applied with full rigour show an object to be mid-way in the relevant respect between standard instances of correct and incorrect application of a concept. For B-concepts uncertainty about their applicability cannot however arise in this way, for the primary tests for their applicability will never yield a final result.

With all these last three sources of uncertainty, there is no need for the logical possibility to be realised. For example, colour concepts are inexact for the last reason. But the world might consist of physical objects which were either of one particular shade of blue or of one particular shade of green or of one particular shade of red, but of no objects of any other colours. Then, if language had colour terms defined by standard examples of each shade, there would be no border-line cases for the application of colour concepts.

The range within which, for the reasons which we have considered, it is uncertain which is the exact value of some concept of measurement is, we have stated, the range of imprecision. But some men are more cautious than others, and what is certain for some is only highly probable for others, and hence the range of imprecision is itself often somewhat imprecise. However, to some limited extent scientists have established conventions for measuring the amount of imprecision of a measurement. Thus, if the measurement is imprecise because there are a number of estimates of the true value (obtained on different occasions or by different individuals), then the mean of the estimates is normally taken as the best estimate of the true value, and the range of imprecision often measured by (e.g.) $\{\Sigma(v^2)/n(n-1)\}^{\frac{1}{2}}$ where there are n estimates and $\Sigma(v^2)$ is the sum of the squares of their deviations from the mean. A considerable majority of the estimates will normally be within the interval. In claiming that the true value of the quantity measured is within that interval, the scientist thus makes precise, formalises, the criterion of ordinary use that in so far as a considerable majority of judgments about what happened coincide, they are to be accepted. A considerable majority of estimates coincide in this respect, that they agree that the true value lies within that interval. Yet of course, men may differ about what counts as a 'considerable' majority, and a wider range than that of the standard error is sometimes adopted by more cautious scientists as marking the imprecision of the measurement. Conventions are very loose in this field.

These then are the various ways in which uncertainty about the applicability of a term may arise, and we have seen how in various cases this may

be due to its vagueness, to the inexactness of the concept which it denotes, or to practical difficulties. Frequently, of course, uncertainty may arise from more than one cause. Uncertainty about the applicability of colour terms is an obvious example. As we have seen, some people have a wider colour vocabulary than others, and so they use colour terms in different ways. Yet even when criteria for the application of colour terms are standardised, observers may make their judgments in bad light. Even when they make them in good light they may differ in their judgments, or agree in judging an object to be a border-line case for the application of a colour term. In these ways its vagueness, practical difficulties, and the inexactness of the concept may all contribute to uncertainty about the applicability of a colour term.

Uncertainty about the answer to some question may often arise from uncertainty about the applicability of two or more concepts. Thus uncertainty about the exact length of a river will be due both to uncertainty about what is and what is not a part of the river and to uncertainty about the exact length of the river, even given its boundaries. Standard examples of stretches of water that are and that are not parts of river can be provided, but near its mouth the stretches of water are borderline cases for being parts of the river, and remain so when primary tests have been fully applied. In Körner's terminology a river is, for this reason, an 'indefinite' object (1966, p. 30). Yet even with the exact boundaries it would be uncertain what is the exact length of the river, because the primary tests for 'exact length' cannot be fully applied.

3

Our second question is whether this uncertainty in the application of terms is a necessary feature of any language in any world, or whether it is merely a feature of our language in our world. We have seen how for many terms there must be logically possible, and for some terms of exact measurement actual, objects, about which we are uncertain whether the term applies to them. Our question is now—must we have terms of this kind in our language, if it is to be a language for describing the physical world? Let us ask in turn of each of the sources of uncertainty whether it must operate in any language in any world.

First, then, what of vagueness of terms? Is it a necessary feature of language that different speakers of the language should use some of the terms in different ways from each other? I cannot see that it is. After all, might not there be a language spoken only by a small group of people, say members of one family on a desert island, who had all learned it together and only talked when together? Then whatever rules had been given for the

use of words, whatever standard examples of objects or states having some characteristic had been produced, would have been known to all. I admit that this is an extreme case. Normally what we call the same language is spoken by different groups who over long periods of time have little contact with each other and who begin to use words in different ways. But it seems in no way necessary that this should be so. The speakers of a language might always use all words in the same ways. True, they would be unlikely in that case to make much progress in their understanding of the world; for such progress normally involves using words in new ways to describe new aspects of things. But then they might just not make such progress.

Now we noted that when the users of language have the same primary tests for the applicability of a term and so use it to denote the same concept, uncertainty about its applicability to some object may remain either because of the inexactness of the concept or because of practical difficulty. Uncertainty due to practical difficulty is temporary, removable uncertainty; uncertainty due to the inexactness of a concept is not. The question therefore remains whether we are obliged to have inexact concepts for describing the world. We will consider this question asking of all the ways in which uncertainty about the applicability of concepts arises whether it is necessary that concepts inexact for these various reasons be used in our description of the world.

First, investigators may be uncertain about the applicability of concepts because they use secondary tests. This uncertainty arises merely from practical difficulty. Second, they may be uncertain because there are different primary tests for the applicability of a concept. If there are different primary tests, so that half of the tests may be satisfied and the other half not, the concept will be inexact. Is it necessary that a language have concepts inexact for this reason? Superficially it may appear that it is not, that if there are different primary tests for the applicability of a concept, it can be replaced by distinct concepts, one corresponding to each primary test.

Mellor (1966, p. 347), however, claims that 'an explainable concept of experience, in short, a scientific concept, must enter into a number of firmly accepted general statements connecting it with other such concepts' (he terms a concept 'explainable' if it 'provides material for theoretical explanation'). He admits (p. 348) that a term may occur only in one sentence stating a 'firmly accepted general statement' but claims that in that case the sentence will merely provide a definition of the term by some other term, and these two terms will then denote the same concept. No distinct concept will have been introduced. Mellor's argument seems fallacious.

A term may be defined by a conjunction or disjunction or other logical operation on other terms. Those terms may denote previously held concepts, but the new term will denote a concept distinct from any previously held. Mellor might in defence claim that though it be possible that a concept enter into only one 'firmly accepted' general statement, such a concept would not be a very useful one for science. That seems plausible enough, but for the fact that Mellor understands (1966, p. 349) by a 'firmly accepted' statement an analytic statement.

Since on his understanding of what is a distinct statement, the 'firmly accepted' general statements will be 'logically independent', it is then logically possible that there be an object which, given the truth of one statement, is an instance of the concept, and, given the truth of another statement, is not. Hence there will always be the possibility of border-line cases. Thus the laws might be $(x)(fx \equiv gx)$ and $(x)(fx \equiv jx)$. Since these laws are 'logically independent' an object could be g and $\sim j$. Such an object would then be a border-line case for being f in virtue of ' f ' having two distinct syntactic definitions.

Yet on Mellor's understanding of 'firmly accepted', I do not find it plausible to suppose that any concept must enter into more than one firmly accepted general statement. A term denoting some concept might be defined by one general statement, which, as a definition, was analytic, yet be useful because it occurred in a number of other synthetic general statements. Certainly, as Mellor urges (1965, p. 109) any sentence stating a law 'can be made to provide an alternative operational definition' of terms occurring in it. But the point is that it need not. One sentence may provide the only definition of some term denoting a distinct concept. I conclude that there is no logical necessity for any language to show conceptual inexactness for the reason which Mellor gives.

Thirdly, is it necessary that a primary test give unclear results? Let us consider the various ways in which a primary test may give unclear results. First, is it necessary that the primary test be not applied with full rigour? Primary tests for the application of any A-concept can be applied with full rigour, but primary tests for any B-concept cannot be applied with full rigour. The effect of ever more rigorous applications of a primary test, we saw, was continually to narrow the range of objects, actual or logically possible, about which it was uncertain whether or not the concept applied to them. Yet for any degree of rigour you like to choose, it is always logically possible that an object remain within the range of uncertainty of applicability of any B-concept, when a more rigorous test would remove the uncertainty. Consequently, because of the impossibility of applying them with full rigour, B-concepts are necessarily inexact.

Do we need B-concepts to describe the world? Why not replace our usual concepts of length, mass, etc., by concepts such that primary tests can be applied with complete rigour? For instance the test whether or not two objects were the same length would be whether they looked as if they coincided to most observers in normal light. The test whether or not two objects weighed the same would be whether or not most observers judged that they balanced each other on the pans of an ordinary scale. It is an interesting question whether or not the use of such concepts would enable us to give as full a description of the world as use of normal B-concepts does; but whatever the answer to this question, the new concepts would also be necessarily inexact—yet for different reasons, and this we shall see shortly.

Let us, however, suppose for the moment that we do use B-concepts of length, mass, etc. for our description of the world. We noted that for many of these it was possible that at a certain degree of rigour of application of tests, no actual objects would remain within the range of uncertainty. Yet given that there are material objects, as there must be in a physical world, then they will have a length and mass, and given that there are events, as there must be in a physical world, they will last some interval of time. A full description of the physical world must state exact lengths, masses and temporal intervals. Yet, as we saw, if these are measured on a dense scale there will be a necessary imprecision in such measurements.

Do we need dense scales of measurement for such properties? Clearly we do if, were we to use any non-dense scale S to measure some property, then there would be another narrower non-dense scale S' (narrower in the sense that each value of S' is a value of S , but not vice versa), such that the values of the property pertaining to some objects would differ on the two scales. Thus suppose we measured the length of objects in inches. It might be the case that all objects were an approximate number of inches long, so that it was clear which value on the inch scale ought to be attributed to each object. Nevertheless if we were to measure instead in tenths of inches, the length on this scale of some objects might differ from their length on the inch scale. In that case the tenth-of-an-inch scale would be better than the inch scale for measuring the exact length of objects. If for any non-dense scale S , a better scale S' could be found, clearly we need a dense scale to make our measurements, for such a scale will allow us to make more accurate measurements than any scale with a non-dense set of values (although the measurements will still not be perfectly accurate).

Yet if we found some non-dense scale S for measuring a property, and had good reason to believe that no narrower non-dense scale would give different results, we would not need a dense scale. What kind of reason

could we have for such a belief? Our reasons could be more or less direct. Most directly, we could find that as we increased the narrowness of our scales and applied our tests to objects more rigorously to discriminate between closer values of the quality being measured, measurements on all tested scales narrower than a certain scale were the same as measurements on that scale. Less directly, it might be a consequence of a very well established fundamental theory that some property could only take a discrete set of values. Direct evidence that various physical properties could be measured adequately by a scale of a certain narrowness has been provided since 1885 when Balmer showed that the frequencies of radiation of hydrogen were discrete. Subsequently Quantum Theory became well established and it is a consequence of Quantum Theory that certain physical properties take only discrete values. Hence I conclude that empirical discoveries can give us reason for replacing dense by non-dense scales of measurement. Yet until we have grounds for choosing a non-dense scale, we ought to use a dense scale of measurement, for otherwise we restrict in advance the information which can be yielded by making a measurement. Yet if we do use a non-dense scale of measurement, it is logically possible that we can find out the exact value of a measurable property, that we can abolish imprecision.

The second possible source of unclarity in the result of a primary test is that observers may disagree about what the apparent results are. Now if this uncertainty is due to the practical difficulty that our observers are untypical, it can be overcome. If, however, typical observers disagree, then the uncertainty about application is due to the inexactness of the concept. Yet if a concept is a concept which applies to the physical world, then different observers must logically be in a position to challenge and check each other's provisional judgments. Whether or not a concept applies depends on whether or not it appears to most observers to apply. For if a concept is such that one observer is of logical necessity in a far better position than any other observer to judge whether or not it applies, it is a concept which describes his mental state and not the physical world. Now observers make their provisional judgments about whether concepts apply to a physical object in the end by comparing and contrasting them with standard instances and non-instances of their application. There is then always the logical possibility that the provisional judgments of observers may conflict so that they are unable to reach a final judgment. It is necessary for the existence of a language for describing the physical world that there be general agreement about the application of terms, but the existence of such agreement is compatible with occasional disagreement. However, as we saw earlier, it seems in no way logically necessary that

observers ever in fact disagree about the application of any concept to any actual object.

Thirdly the unclarity may arise because of the different results of different applications of the test, e.g. with different instruments or on different occasions. Now if this is due to practical difficulties it can be overcome by using better instruments. But if it cannot be overcome by using better instruments, it will be due to the inexactness of the concept. While there is no necessity that actual instruments ever yield different results from each other or yield different results on different occasions, it seems necessary that our concepts be such that this is logically possible. This is because if we are to describe our world we must suppose that objects behave in a regular way. Only if we suppose that they do can we predict behaviour or even describe anything beyond our immediate perceptual field. The supposition that an object behaves in a regular way is the supposition that it continues to possess over time a certain property. If we are to describe the world beyond our perceptual field we must ascribe to objects such properties enduring for a period of time as having a certain mass, losing mass at a constant rate, travelling with a certain acceleration, or emitting light with a certain wave-length. The best evidence from a single application of a test which we can have for ascribing the property will be the evidence of the result of a test made during the period in question. Yet since the property is an enduring one, the result of any other application of the test during the period will be equally relevant. Yet these applications could sometimes give different results. Then it would be unclear whether the object had the property.

Fourthly the unclarity may arise because the final result of all applications of the primary tests is that the object is a border-line case for the application of the concept. Now this situation cannot arise with B-concepts, for with them we cannot get a final result for our primary tests. But it can and does arise with A-concepts, as we saw with the example of colour terms, and is to be attributed to the inexactness of these concepts. Körner in particular, as we saw, (1966, p. 26) has analysed uncertainty arising in this way and claims that 'many resemblance classes are inexact' (he understands by a 'resemblance class', a class, the membership of which is determined by the similarity of objects in some respect to standard members or non-members of the class). He does not, however, appear to claim that all resemblance classes are inexact, nor that all classifications of physical objects are classifications into resemblance classes. We must therefore investigate independently to what extent Körner's kind of inexactness is a necessary feature of language.

Concepts are, we have seen, ultimately linked to the world by standard

instances and standard non-instances of their application, objects which are standardly f and objects which are $\sim f$. Now for any two objects one of which is $\sim f$ and the other of which is f a series of objects could, it is logically possible, be constructed intermediate in all relevant respects gradually more and more like the one and less and less like the other. Objects could be constructed intermediate in size, shape, colour, smell, taste and feel, that is in all the properties which our senses can detect. Of all such properties it makes sense to suppose that degrees intermediate between any two distinct degrees are possible. Therefore a series of intermediate cases, of objects more and more like standard f -objects and less and less like standard $\sim f$ -objects can be constructed. So whatever our standard of an object mid-way between the standard f and standard $\sim f$ objects, the existence of such an object and hence the existence of a border-line case for the application of the concept f is logically possible. If the concept is a B-concept no actual object could be shown conclusively to be such a border-line case, for further tests could be applied to show to which group the object was really more similar. But if a concept is an A-concept tests can be applied with perfect rigour and an object could be shown conclusively to be a border-line case. I conclude that, for different reasons, both A-concepts and B-concepts are necessarily inexact. We might be able to substitute B-concepts for A-concepts or vice-versa but the inexactness would remain.

Waismann (1945) has adduced a further reason other than those which I have so far considered for holding that inexactness is a necessary feature of many empirical concepts and this is what he calls their 'open texture'. He claims that 'no definition of an empirical term will cover all the possibilities' (1945, p. 123). Any attempt to lay down the conditions of use of some term will never be able to cater for every possibility, because it is always possible that I might have some totally new experience 'such as at present I cannot even imagine' (1945, p. 124) or some new discovery may affect our whole interpretation of facts. Now Waismann does not claim that inexactness is a feature of all empirical concepts, yet he seems (p. 121) to claim that it is an essential feature of most. The grounds, however, which Waismann gives for his claim, have a peculiarly elusive character. If he claims that some concepts must be inexact because I might have some new experience 'such as at present I cannot even imagine', it is no good my attempting to refute this by claiming of a particular concept that no such experience is possible. He would say that I claim this because I cannot imagine the possible experience. Any attempt of mine by pointing to kinds of possible experience to show that nothing could be a border-line case for some concept would be irrelevant, because Waismann is talking

of experiences that at present we cannot even imagine. Certainly some past concepts have seemed exact enough, until men have had some hitherto unimagined experiences, and then some objects have turned out to be border-line cases for the application of concepts. But other concepts might not have been afflicted in this way, and I cannot see any argument to show that any concepts must be. The same argument applies to Waismann's suggested new discovery which would affect our whole interpretation of facts. I cannot see why, for any concept, there must be a logically possible discovery which would show—what was previously not realised—that certain objects are border-line cases for the application of the concept.

Nevertheless for the various reasons which I have given previously I conclude that for each of our empirical concepts there could, it is logically possible, be an object about which we would be uncertain whether or not the concept applied to it, even when we had applied tests for its application with as much rigour as we chose. Our concepts are therefore inexact, and we have now shown that this inexactness cannot be removed from a language for describing the physical world. Yet, we also saw that, with the exception of concepts of exact measurement on a dense scale, it is in no way necessary that the logical possibility be realised. There is no reason to suppose that, with the stated exception, we could not choose concepts for description of a world, by means of which we could describe all recognisable aspects of that world, and for the application of which there were not in fact any border-line cases. There might, for any *f*, be no objects about which it was uncertain if they were *f* for any of the reasons which I have delineated. However, if we use concepts of exact measurement on a dense scale, there will be a necessary imprecision in any measurement. We could, however, dispense with such concepts in our description of the world, if we replaced our B-concepts of measurement by A-concepts, and let the judgment of men unaided by instruments be the final criterion of 'same length', 'same temporal interval', etc. It is, however, open to question whether we should regard description of the world by means of terms denoting such concepts as giving as full a description of the world as description by terms denoting our current B-concepts. However, we saw that, even if we use B-concepts, in certain empirical circumstances there would be no need for us to make our measurements on a dense scale, and we saw evidence that these circumstances in fact hold, at any rate for measurements of some properties. If the length of objects, period of events, energy of systems come in quanta, we do not need dense scales to measure them. In that case there need be no uncertainty in the application of concepts to objects in giving as full a description as we like of some physical world.

University of Hull

REFERENCES

- ALSTON, William P. (1964). *Philosophy of Language*. N. J.: Englewood Cliffs. Chapter 5.
- BLACK, M. (1949). Vagueness: an exercise in logical analysis. In *Language and Philosophy*. Ithaca, New York.
- KÖRNER, S. (1966). *Experience and Theory*. London.
- KUMAR, Dharmendra (1967). Logic and inexact predicates. *Br. J. Phil. Sci.* 18, 211-22.
- MELLOR, D. H. (1965). Experimental error and deducibility. *Phil. Sci.* 32, 105-22.
- MELLOR, D. H. (1966). Inexactness and explanation. *Phil. Sci.* 33, 345-59.
- MELLOR, D. H. (1967). Imprecision and explanation. *Phil. Sci.* 34, 1-9.
- WAISMANN, F. (1945). Verifiability. *Proc. Aristotelian Soc. Suppl. Vol. 19*, 119-50. Reprinted in A. G. N. Flew. Ed. (1951, pp. 117-44): *Logic and Language* (1st series). Oxford. Page references in the text are to this latter volume.

Action, Causality, and Teleology

by RUTH MACKLIN

I

A significant dispute in the philosophy of action centres on the question of the appropriateness of selecting certain patterns of explanation, rather than others, for the purpose of explaining human behaviour or action. A related problem deals with the question of the applicability of certain conceptual categories which are usually held to be in opposition to one another: *causes* of action versus *reasons* for acting; *causality* versus *purpose*; *causal explanations* versus *teleological explanations*. In this paper I shall focus on the second and third pairs of opposing conceptual categories. I shall argue that not only are purposive explanations compatible with causal explanations, but also that the former—when suitably recast—can plausibly be construed as a species of the latter.

The notion of purpose is usually taken to be crucial for characterising the concept of human action, as well as for understanding and explaining particular actions. As a minimum explication of the concept of purpose, we may say that the concept expresses the notion of goal-directedness and bears the same sense as the locutions 'in order to', or 'for the sake of'. When we say: 'X's purpose in doing *A* was to do *B*', this expression can be replaced by the expressions 'X did *A* in order to do *B*' or 'X did *A* for the sake of doing *B*'. Of course, this minimum account does not go far in explicating the notion of purpose, but merely shows the replaceability of expressions containing the term 'purpose' by other terms which express the same concept. Another concept related to purpose is that of *means and ends*. And all of the notions mentioned comprise what is meant by the term 'teleological'. There is no doubt that human beings do act purposively, that they do things for the sake of other things, that they usually have goals or ends in view and that their behaviour is almost always goal-directed in some way. The question arises, however, whether the fact that actions are essentially purposive or goal-directed entails that explanations of human action need be teleological in form. A second question which immediately follows this one is related to the broader issue of conceptual frameworks mentioned above. This question is whether or not the employment of teleological categories precludes the

Received 8 August 1968

employment of causal categories in explanation of action. In this paper, arguments are presented to the effect that both of these questions are to be answered in the negative, although it might be maintained that the first of these questions can be answered positively without entailing a positive answer to the second.

Among those writers who deny the applicability of causal categories to explanation of human action, there are a number of different viewpoints maintained and it is not always clear precisely in what way the notion of cause is being employed. In section 2 below, I shall propose a sense of 'cause' which, although it may depart from the standard Humean or 'mechanistic' sense of 'cause', is nevertheless a relevant sense of 'cause' for explanation in the natural sciences as well as for an account of human action. Accordingly, I am not so much interested in whether or not the position I adopt here is compatible with that of some or all of the anti-causal theorists. Rather, the point is to demonstrate the way in which explanations of human action—while incorporating the purposive or goal-directed features of human action—can be construed as causal explanations, in some sense of 'cause' that is relevant and appropriate both for the natural sciences and for the science of human behaviour.

Before proceeding to the main arguments of the paper, some preliminary remarks need to be made by way of setting the stage for the position advanced. A significant group of writers¹ is committed to the view that causal patterns of explanation are in general inappropriate for human action. Several interconnected theses serve to characterise the position of these writers. In some cases these theses have been stated explicitly, in other cases they have been tacitly assumed. In general, however, whether the theses operate as assumptions or as explicit commitments, their acceptance is crucial to the view that we cannot provide causal explanations of human action. It may be that no one writer is committed to all of the following theses, but it seems that a reconstruction of a full statement of the view in question is consistent with the acceptance of all of these commitments.

- (i) Causal explanations are (always, generally, usually) mechanistic explanations.
- (ii) In order to provide causal explanations of human behaviour, psychological terms (e.g. 'motive', 'intention', 'want', 'desire') must first be reduced to physiological terms, electro-chemical terms, etc. Such a reduction would be necessary for obtaining the general laws required for any causal explanation.
- (iii) In explanations of action, psychological terms cannot be used to

¹ Representative writers in this group are A. I. Melden, R. S. Peters, Peter Winch, Anthony Kenny, D. W. Hamlyn, and Stuart Hampshire.

denote *causes* of action, although they may appropriately be cited as reasons for action.

- (a) Motives, desires, intentions, wants, etc., are not of the right logical type to be causes; hence, it is a conceptual error to cite them as causes.
- (b) Intentions, wants, motives, etc., are logically, not contingently, linked to actions; hence, they cannot be causes of actions.
- (iv) There is a clear distinction which can be made between what a person *does* and what *happens* to him.
- (v) Human behaviour is rule-following behaviour. Any view of action which fails to take this feature into account cannot be an adequate theory of action.

Although I shall not argue here specifically against any of the above five theses,¹ the account which follows is contingent upon the admission that one need not accept these theses uncritically and that failure to adhere to them opens the way for a radically different theory of action from that espoused by the writers in question. Accordingly, although the reader may be inclined to demur at some of my claims below, I ask only that he approach the issues with a principle of charity, recognising that sound arguments against the above five theses can be (and have been) given.²

2

While it is beyond the scope of this paper to provide a comprehensive account of causality or to construct in detail a notion of cause which is non-mechanistic or non-Humean, a few points should be made in this connection. First, there seems to be no reason to hold that any explanation which essentially involves causal laws must be a mechanistic account (i.e. involving mechanical laws or mechanistic predicates). It should be emphasised that the relevant sense of 'cause' for any explanation (whether of human behaviour or in the natural sciences) must include the notion of generality. There needs to be subsumption under some law, although we may not now have the law available. Thus we can refer to 'causes' or to 'causal factors' on the assumption that the condition of generality will be met. This notion of cause can be characterised as follows (Gorovitz (1965), pp. 706-7):

¹ I have argued against thesis (iv) in 'Doing and Happening', *Review of Metaphysics* (forthcoming) and against each of the remaining theses in my unpublished dissertation, *Theory of Action*. The arguments given are directed as well against the general position that it is a conceptual error or otherwise inappropriate to attempt to provide causal explanations of human action.

² For further examples of arguments attacking one or more of these theses, see Donald Davidson (1963), W. D. Gean (1965), May Brodbeck (1963), J. J. C. Smart (1964).

The claim that some factor *c* is the cause of some event *e* . . . entails the claim that *c* is causally relevant to *e*. And this entails that there is some law connecting *c*-like events with *e*-like events (though not necessarily or even usually under the descriptions we use when we cite *c* as the cause or ask for the cause of *e*). Thus, if we were to attempt to justify the claim that *c* satisfies the conditions—and hence can be the cause of *e*—we might have to provide a law, or evidence that there is one, connecting *c*-like events with *e*-like events.

Second, the undesirable features of a Humean account of cause, in this context, seem to be those centring around the epistemological issue of knowledge of the constant conjunction of the putative cause and effect. But that feature of a Humean account of cause requiring knowledge of the constant conjunction of *c* and *e* is a feature which is often absent from explanation in the natural sciences. So if a non-Humean, non-mechanistic notion of cause could be provided, it might be shown to be a relevant sense of 'cause' both for explanation in natural science and for human action.

There have been a number of recent attempts to provide a non-Humean, non-mechanistic sense of 'cause'. Miss Anscombe, in her book *Intention*, attempts to provide an account of at least one non-Humean sense of 'cause' with her notion of 'mental cause' (pp. 13 ff.). In this sense of 'cause', 'the subject is able to give the cause of a thought or feeling or bodily movement in the same kind of way as he is able to state the place of his pain or the position of his limbs' (p. 10). The point here is that in the case of mental causation the subject is aware without observation that *x* causes *y*. Miss Anscombe's account of 'things known without observation' is obscure and difficult, but we may bypass these problems here. It is not my aim to explicate in detail Miss Anscombe's notion of 'mental cause', but only to note that it does refer to a recognisable human phenomenon and does not fit the Humean conception of cause, as she remarks (p. 16):

Note that this sort of causality or sense of 'causality' is so far from accommodating itself to Hume's explanations that people who believe that Hume pretty well dealt with the topic of causality would entirely leave it out of their calculations; if their attention were drawn to it they might insist that the word 'cause' was inappropriate or was quite equivocal. Or conceivably they might try to give a Humean account of the matter as far as concerned the outside observer's recognition of the cause; but hardly for the patient's.

Another account which is similar in some respects to Miss Anscombe's and draws on some of her remarks is that of John Yolton in an article entitled 'Agent Causality'. Yolton agrees with Miss Anscombe that the notion of cause has been left in a state of confusion, and he claims that this is 'a consequence of the analysis of cause as a uniform and ordered sequence, the analysis initiated by Hume and refined by Kant. . . . [It] made our knowledge of cause the result of repeated observations, of

induction' (Yolton (1966) p. 22). Yolton's positive suggestions are rather obscure, but they point in the direction of a non-Humean notion of cause which at the same time attempts to avoid any implications of mechanism. He writes (p. 22):

The notion of cause as correlation, as ordered sequence, is clearly inadequate when we seek to formulate our thinking about action. . . . To say '*X* causes *Y*' usually involves the thought of energy expended, of chemical or electrical processes, of motion and impact. There may be no need to retain these notions of cause in our sophisticated philosophy of nature or of science. We need them in our philosophy of man. Yet, paramechanics threatens if we try to apply these pushings and impacts, even these electrical and chemical changes, to thought and action.

Although Yolton's account is incomplete and somewhat obscure, it does attempt to seek a notion of cause which avoids the implications of mechanism. In addition, this notion of cause is 'non-Humean' and so it applies to the area of human action in a way which the Humean notion of cause fails to do. Yolton has more to say along the same lines in his article (1966a) 'My Hand Goes Out to You', where he maintains that in many ways the operative concept in action is cause. Here again he denies that the Humean notion of cause is applicable to action (p. 149):

If we stay with Hume's dictum about not being able to say, prior to experience of the conjunction of *C* and *E*, anything about *E* on the basis of *C* alone, we must exclude cause from actions. . . . For one of the features which distinguishes an action from mere bodily movement is the knowledge of what I am going to do. Since I cannot be a moral person, i.e. responsible for what I do, unless I am the agent of my action, we must quickly conclude that Hume's dictum does not hold for instances of mental causation.

Yolton's view departs from the Humean notion of cause in the following way. If the Humean conception requires that one have experience of the conjunction of *C* and *E* before ascertaining that *C* is the cause of *E*, then Yolton's notion of mental cause fails to satisfy this Humean requirement. Nevertheless, the ways in which this conception of mental cause enters into much human behaviour are evident. As an example of a mental cause, we may cite the appearance of a face at the window, as a result of which I become frightened and telephone my neighbour. The appearance of the face at the window causes my fright and my subsequent (intentional) action of telephoning my neighbour. My knowledge of this fact (i.e. that I am frightened because of the appearance of the face) is non-inductive and is not based on prior experiences of faces at windows conjoined with my being frightened and telephoning my neighbour. In fact, I may never have experienced any conjunction which was similar (in relevant respects) to this one. It appears to be Yolton's point that many

cases of human behaviour are of this sort, where we do not have prior experience of the conjunction of *C* and *E* and yet we know, on the basis of *C* alone, that *C* is the cause of *E*. Moreover, it is absurd to expect that we need to have a causal law of the form: Whenever a person sees a face at a window and is frightened by that face, he telephones his neighbour. If, indeed, we can meaningfully and plausibly speak of the causes of human action, it does not follow that the causal antecedents will be subsumable under laws whose descriptions contain the kinds of terms appearing in the above-mentioned example.

In order to carry the analysis beyond that offered by Anscombe and Yolton, several points need to be noted. For causal explanation to be appropriate for human action, at least the following minimal conditions must be satisfied: (1) there must be a sense of 'cause' which does not involve the implications of mechanism; (2) there must be a sense of 'cause' which avoids certain features of the Humean account (as noted by Yolton and others)¹; (3) this sense of 'cause' is the relevant sense for explanations of human action; and (4) the condition of generality (i.e. subsumability under some law) must be met. As to a further claim which Yolton makes—that on some notion of cause other than a mechanistic or Humean one, motives, intentions, decisions, volitions are examples of mental causes—the arguments on both sides need to be examined. But certainly, if we accept a sense of 'cause' which satisfies conditions (1)-(4) (whatever the details of such an account might be), then we are in a better position to admit motives, intentions, decisions, etc., as candidates for causes of human action than we were formerly. However, in order to meet fully the objections of the opponents of causal explanations of action, we need to show that the features of this notion of cause are features found in causal explanations in the natural sciences. If it can be shown that the non-mechanistic, non-Humean sense of 'cause' is operative in significant cases of explanation in natural science, this result would partially undermine the claim that the pattern of explanation is essentially different in the natural sciences and in the science of human behaviour.

There seem to be several kinds of cases in physical science where the notion of cause departs from the Humean and/or mechanistic conception.

¹ The rejection of still another Humean feature in the notion of cause can be found in R. Hancock (1962). This feature consists in the view that the observation of past sequences of similar events is necessary for the imputation of causality. Hancock, arguing against the view that human actions cannot admit of causal explanations, makes the following points. 'At most, one might show empirically that the behavior of persons is less consistent than the behavior of physical objects—in the sense that different persons or the same person at different times react very differently to the same threat, command, or the like. But this would not prove that in interpersonal causation the same cause need not have the same result in the same circumstances' (p. 371).

First, there are cases in which the antecedent conditions are known to be the cause without prior observations of this sequence. The particular phenomena—under this description—were never before observed to act in this particular way. The obvious example is the first explosion of an atomic bomb. Similarly, in cases of chemical combination, the combination of these particular chemicals, followed by this reaction, may never before have been observed. It can be known that combining the chemicals resulted in the new compound, and, hence, a causal explanation of the phenomena might be given even though this particular cause-and-effect sequence was never before observed. (Of course, a complete causal explanation will undoubtedly include mention of the micro-structure of the chemicals, and so the description under which the event is explained need not be couched solely in terms of observables.)

Other examples are those of a more everyday sort. An individual, having had no previous experience of a particular sequence, knows that a particular set of antecedent conditions led to a particular result. For example, if Jones puts his hand in hydrochloric acid, thinking that the liquid is water, he knows the cause of the resulting pain without any prior observations of a similar sequence (i.e. he may never have immersed his hand in any substance which was not water, and yet which resembles water). He may not be able correctly to describe the substance as hydrochloric acid in this case; yet he knows that immersing his hand in that substance—whatever the correct description may be—caused the pain. Cases of this sort could be multiplied, lending support to Donald Davidson's claim (1963, p. 698) that 'we are usually far more certain of a singular causal connection than we are of any causal law governing the case'. It is in precisely these cases—cases of singular causal connections—that there need not have been repeated observations of a constant conjunction in order to know that *A* caused *B*. This seems to be the typical case in human action, and although this type of case is less common in physical science, it seems to exist nonetheless. Davidson makes an additional point which is important for noticing the analogy between at least some cases in physical science and the typical situation in human behaviour (p. 697):

Ignorance of competent predictive laws does not inhibit valid causal explanation, or few causal explanations could be made. I am certain the window broke because it was struck by a rock—I saw it all happen; but I am not (is anyone?) in command of laws on the basis of which I can predict what blows will break which windows. A generalization like 'Windows are fragile, and fragile things tend to break when struck hard enough, other conditions being right' is not a predictive law in the rough—the predictive law, if we had it, would be quantitative and would use very different concepts. The generalization, like our generalizations

about behavior, serves a different function: it provides evidence for the existence of a causal law covering the case at hand.

Among the writers who reject the Humean notion of cause, none denies that at least some generalisation underlies a singular causal judgment. The point here is simply that the causal model which some writers claim obtains for all physical phenomena, and which they deny application to human phenomena, is by no means the only (or perhaps not even the prevalent) causal model. As R. Hancock writes (1962, p. 376):

Where the cause is a familiar human action altering the normal course of events and the result follows immediately, it can be argued that the sequence is recognizably a causal one even though similar cases have never been observed. One can surely see in a single case that, for example, hammering glass causes it to break, without ever having experienced this sequence before: the use of causal language no doubt presupposes some kind of regularity in experience, but the same is true of the use of causal language in interpersonal causation. . . . I conclude, therefore, that while it is true many cases of physical causation cannot be recognized as such apart from past observation of the same sequence, this is not in general true of physical causation and does not prove a difference in meaning between causal language in interpersonal and physical causation.

It should be noted, finally, that explanations of physical events, as well as of human actions, may be exceedingly complex. The matter is oversimplified by an attempt to search for efficient causes or for *the* cause of any event or action. It is certainly the case in science that deep explanations frequently include terms which are embedded in complex theories, and that such terms sometimes denote theoretical entities. It is plausible to hold that theoretical terms are to be found, as well, in descriptions of psychological phenomena, and so many explanations of human action will include reference to unobservables or to micro-phenomena. In these cases of complex scientific explanations involving theoretical terms, we do not ascertain the causes simply by noting a constant conjunction between observable antecedent conditions and subsequent effects. For example, in biological phenomena, the cause (or one of the causes) of aberrant animal behaviour may be attributed to a mutation or to some other genetic factor. The mutation may have resulted because of high radiation in the environment. A complete explanation of the phenomenon will not only mention the radiation as the cause of the aberrant behaviour, but will also refer to a postulated change in the genetic make-up of the animal. The appeal to unobservables or to micro-states or processes is made on the basis of biological theory, in particular, genetic laws. In such cases, although the particular conjunction of macro-events had never before been observed, a causal explanation can be constructed nonetheless. If it is plausible to construe a number of psychological terms as theoretical

terms, as I shall later suggest, then explanations of human action can be shown to have strong similarities to the type of biological explanation just cited. It would seem to follow, then, that either a non-mechanistic or non-Humean notion of cause is appropriate for explanations of *both* natural events and human actions, or else many explanations in physical science are to be deemed non-causal.

In sum, the account in this section has tried to show that there is a notion of cause which does not entail mechanism and that this conception of cause is germane to human behaviour. Although none of the accounts discussed above is fully adequate for providing the requisite sense of 'cause', the several attempts are worth noting because they point in directions which might be fruitful and which require more philosophical attention. Moreover, it has emerged that in the physical sciences there are counter-examples to the standard Humean account according to which our knowledge of causes is always held to be inductive knowledge. As Yolton notes (1966a, p. 150):

If we free ourselves from some of the Humean notions of cause, as well as from the mechanical notions of cause as pushes, we may free ourselves for talking of mental causes in the important sense of mental events which control behaviour.

3

In light of the preceding discussion, we may now go on to examine the role that the notion of purpose plays in explaining actions within a causal framework. Let us consider, first, the view that causality and purpose are incompatible conceptual categories. One of the theses of Richard Taylor's book, *Action and Purpose*, is that the concepts of *purpose* and of *cause and effect* are basic categories, irreducible to one another.

Causation is an ultimate philosophical category, or one that has not been, and evidently cannot be, analyzed in terms of any other philosophical ideas. . . . I shall defend the claim that *purpose* is another such ultimate philosophical category—that it too is a concept that has not been and evidently cannot be analyzed or understood in terms either of causation or of any other philosophical ideas. (Taylor, R. (1966) p. 12.)

To deal with the matter in Taylor's own terms, one might reply that the question whether purpose is an irreducible *philosophical* category is ill-formed. That is, one might reply that purpose is an irreducible philosophical category in one sense, but not in another—a reply which will be amplified below. Taylor later goes on to note (p. 142) that an explanation 'in terms of one's purpose, aim or goal, is of an entirely different type from an explanation in terms of causes and can in no way be reduced to it'. But this result in no way follows from the remarks in the preceding

quotation. Perhaps part of the problem lies in the vague conception of what is to count as a teleological explanation.

What could it mean to say that purpose is a basic philosophical category? It is certainly true that talk about purpose is not *synonymous* with talk about cause and effect. Moreover, if an agent says: 'I did *A* in order to do *B*' he is almost always not *thinking* of his action in terms of cause and effect. In this connection it is appropriate to cite a remark by David L. Perry (1965, p. 247), made in a slightly different context:

Recognition that as a deliberating agent one does not seek causes of his decision might prompt the erroneous conclusion that an agent's rational processes cannot themselves be natural, psychological causes of his decision. It is well to emphasize that the categories of the agent are not those of the spectator, but it is wrong to conclude that the spectator's categories cannot be applied to the agent's *employment* of his categories.

Perry's remarks about an agent's decisions can equally well be applied to an agent's actions. It seems plausible to hold that an agent's wanting *B* or having a pro-attitude toward getting *B*, along with his *belief* that doing *A* is a means of obtaining *B*, count as causal factors in an explanation of his action, *A*, under some description of that action. Here the notion of purpose is not being *eliminated*, but rather, is incorporated into an explanation of a person's action, an explanation which may be causal in form. Thus, it may be that the means-end relationship, or the 'in order to' locution is irreducible and that that is what might be meant by saying that purpose is a basic or ultimate category. But that result does not seem incompatible with the claim that we can provide causal explanations of human behaviour under some appropriate description of an action. For an agent's forming an intention or making a decision to do *A* in order to get *B*, conjoined with other factors, may be the causally relevant factors in our explanation of his action. Thus, even if Taylor is correct in claiming that purpose is an ultimate philosophical category (in the sense of a conceptual or metaphysical category), it does not follow that all explanations of action need to be teleological in form. By casting purposive elements in terms of the agent's perceptions, knowledge, wants, beliefs, etc., we can consider some purposive *elements* to be basic constituents of an explanation which is similar in *form* to that in the natural sciences. This is not the same thing as reducing purposive or teleological talk to causal talk, but is rather incorporating the former, as constituent elements, into the latter. From the fact (if it is a fact) that a concept is an ultimate philosophical category, it does not follow that such a concept cannot be used as a constituent of a more complex philosophical account which falls into a different 'basic category'.

4

Of course, these remarks are totally beside the point if what is meant by 'teleological explanation' is: 'Any explanation into which enter in any manner whatsoever, terms such as "purpose", "aim", "goal", "means-ends", "in order to", "for the sake of", and so on.' In that case, all or almost all of the explanations we give of human action will be teleological; but that sense of 'teleological explanation' is quite uninteresting and trivialises the issues involved. Moreover, from the adoption of that sense of 'teleological explanation', nothing follows concerning the possibility of such explanations being either causal or non-causal. A better sense of 'teleological explanation' is one which refers to the *form* of the explanation or to the inferential pattern, and not simply to the terms which are employed in the explanation. 'Teleological' is not a term like 'physicalistic' or 'phenomenalistic' which, when conjoined with the term 'explanation', is explicated simply by means of the kind of predicates used in such explanations. A teleological explanation is one in which the laws employed in giving the explanation are teleological in form, but whether such laws are 'irreducibly' teleological is another question. One prominent view is that teleological explanations can be cast in an equivalent non-teleological form:

The difference between a teleological explanation and its equivalent non-teleological formulation is thus comparable to the difference between saying that *Y* is an effect of *X*, and saying that *X* is a cause or condition of *Y*. In brief, the difference is one of selective attention, rather than of asserted content. (Nagel (1961) p. 405).

In opposition to that view, Charles Taylor (1964) discusses the form which the laws of human action or behaviour must take. His view amounts to a denial of Nagel's claim that teleological explanations can be reduced to equivalent nonteleological formulations. Taylor argues that the laws governing action can be shown to be teleological in form on the basis of empirical research by psychologists; hence, explanations of human action into which such laws are incorporated are teleological explanations. He holds, moreover, that such explanations cannot be reduced to non-teleological ones. Now the reduction which Nagel offers is specifically constructed for biological systems and other goal-directed systems of non-organic mechanisms (e.g. thermostats and servo-mechanisms). The question remains open whether such a reduction can be carried out for systems which are not merely goal-directed but are purposive in a stronger sense. Charles Taylor argues in favour of the irreducibility of teleological explanations *in general*, and not simply those pertaining to

purposive human or animal behaviour—but that larger issue will not be tackled here.

It should be noted again that Charles Taylor takes the question of whether the laws governing action are teleological or not, to be an *empirical* question and much of his book is concerned with evidence from the empirical research of psychologists in support of his own claims about teleological laws and explanations. But even if we concede that Taylor is right and Nagel wrong on the issue of the reducibility of teleological explanations, there are still the following difficulties with Taylor's position. He argues against the view, held by those who reject explanation by purpose, 'that the issue can be settled *a priori*, that is, without having to examine the type of explanation which actually applies to living things' (p. 72). Accordingly he holds that such an examination leads to the conclusion that the laws governing action are teleological in form. But it seems that the most that empirical research could show is that human action is essentially purposive or that most actions involve an awareness on the part of the agent of his purposes, aims, goals, intentions, etc. However, this fact alone—namely, the essential purposiveness of human action—does not warrant the conclusion that we cannot give causal explanations of action, as already argued above. In addition, it is difficult to see how empirical research of any sort could inform us of the *form* which the laws of behaviour or the laws governing action should take. The point here is essentially that of Pierre Duhem, namely, the empirical data do not uniquely determine the theory. That point applies as well to the natural sciences as to the social sciences. And in the context of the inquiry which Taylor discusses, the empirical data do not uniquely determine the theory or the form which the laws governing behaviour must take. Thus, even if Taylor can meet Nagel's arguments concerning reducibility, Taylor's own position can be questioned on two counts: first, the point just noted, that the empirical data cannot inform us of the *form* which the laws must take; and second, the earlier point that the purposive elements of human behaviour can themselves be incorporated into causal explanations of that behaviour. If these objections are sound, they go at least part way toward undermining Charles Taylor's attack on the possibility of causal explanation of action.

5

There is, finally, a pragmatic issue relating to action and purpose—an issue that concerns the question of the *adequacy* of purposive explanation. The notion of the adequacy of any explanation depends, in part, on the

questions asked and the purposes which the explanation is meant to serve. Richard Taylor (1966, p. 218) writes:

It seems hardly doubtful, then, that purposeful explanations are often useful, true and adequate in the realm of human behavior. Indeed, it is precisely the kind of explanation that is usually sought in almost every aspect of life.

The question remains whether or not the requirements that an explanation be 'useful, true and adequate' are sufficient for all the purposes we may have. It would seem that there are other conditions for adequacy, such as that of providing systematic interconnections with known theories or that of being a 'deeper' or 'higher level' explanation or one which is more complete than others which might be offered. Taylor's notion of adequacy here seems to be sufficient for some purposes but not for other purposes which are worthwhile. He remarks (p. 218):

We are asking for an explanation that will represent the behavior in question as the means to some end. When such explanations are given, they are sometimes adequate; they set curiosity at rest, they render the behavior inquired about intelligible. They may, in fact they certainly do, differ radically from other types of explanation—from those sought in the physical sciences, for instance—but they do not differ by explaining nothing. The answers they provide are sometimes the only possible answers to the questions that are asked.

But an answer which, at one time in the history of scientific inquiry, is 'the only possible answer to the questions asked', may emerge as barely adequate or even as wrong at some later stage in the development of the science. The adequacy of an explanation is relative to the context of inquiry, the knowledge and purposes of the inquirer, etc. The notion of adequacy of an explanation is not simple and uniquely specifiable but must be relativised to these further factors, especially to avoid closing off avenues of future inquiry. Richard Taylor, like a number of recent writers, holds that an explanation of a bit of behaviour is adequate if it renders that bit of behaviour intelligible. But merely to render an action or piece of behaviour intelligible is not sufficient for the purposes of psychological theory. In constructing and confirming psychological theories, the theorist wants to do more than simply *understand* the behaviour, in the sense in which it is rendered intelligible by a purposive explanation. The theorist may wish to know which other factors determined the agent's action in order to understand better the action-type and the kinds of causally relevant factors which are present. Thus, merely rendering the action intelligible by noting the agent's purpose or the means-end relationship will not be sufficient for the psychological theorist who wishes to know not simply *what* the agent is doing, but also *why* he is performing such and

such a purposive action and what other factors are related to his performing this action. Such factors need not be given in any explanation which is adequate according to Taylor's notion of adequacy, but there seems to be no reason to accept his view of the adequacy of explanations, unless we mean by 'adequate': '*merely* adequate, but not the best that can be offered.'

Ordinarily, though not always, a request for an explanation of human behaviour arises when the behaviour deviates in some way from our normal expectations. A knowledge of social contexts, common practices, and the means-end relationship will in many cases serve to identify the action as being an action of a certain sort and so will inform us of what the agent is doing, in the sense that his behaviour is rendered intelligible. There is a supposition on the part of some writers that once an action is seen to be normal, usual or expected, the question why the agent performed the action does not arise. It is alleged that such a question only arises when the behaviour is in some sense deviant and then it is appropriate to search for the causes of deviant behaviour. But this is surely a *psychological* rather than a *logical* condition being imposed on the propriety of certain types of explanation as opposed to others. Although it is true that the request for an explanation in any given case will be relative to the knowledge of the inquirer, it does not follow that the request for a causal explanation is rendered logically inappropriate simply because the inquirer is familiar with the context or is made aware of the agent's purposes. I have argued earlier that limiting the notion of cause to a mechanistic or Humean conception which can view causes only as antecedent movements is an unwarranted piece of philosophical legislation. Now it can be claimed further that normal or usual events, whether purely physical or instances of human action, also admit of causal explanations, and that the notion of a person acting purposively in performing an action must itself be explicated in terms of that person's beliefs, perceptions, wants, intentions, etc. Whether or not the request for an explanation of this normal, expected behaviour will ever arise is a psychological and not a logical issue. With respect to physical events it is certainly appropriate to offer explanations of expected events in terms of the usual causes and relevant laws: e.g. dissolvings, the falling of bodies, the adaptive behaviour of animals, etc. It would, perhaps, be psychologically surprising but not logically odd for a scientist to request an explanation for an instance of these everyday, well-known occurrences. Similarly, with respect to human behaviour, it would be psychologically surprising for an observer familiar with an agent's purposes and the means-end relationship to request an explanation of an ordinary, non-deviant piece of behaviour. But such explanations are not logically inappropriate and the chief problem seems

to be that of finding or formulating well-confirmed psychological laws under which action-events are subsumable.

6

The arguments in this paper have purported to show that an attempt to construct explanations of human action which are causal need not ignore the fact that human action is goal-directed or purposive. That is, we can incorporate into explanations the features of human behaviour which make it goal-directed in that we can include mention of a person's goals, aims, purposes, intentions, etc., in our description of his thoughts, wants, and beliefs. This is not, in any strict sense, a *reduction* of teleological language since it does not *eliminate* the teleological terms; rather, it places teleological terms in a context in which they can be used in causal explanations of actions. There is, however, a kind of objection which might be raised here and which is argued at some length by Richard Taylor. It is part of Taylor's general thesis that purposive explanations cannot be reduced to causal explanations. But he writes, in addition (p. 222):

A purported causal explanation of behaviour, which represents one's actions as the literal effects of certain inner or subjective states, is nothing but a variant of the theory of volitions all over again.

The arguments against the theory of volitions are by now well-known¹ and Taylor has nothing new to add on this score. In any case, such a result would be pernicious only if psychological terms are taken to denote inner processes or events which are introduced in an *ad hoc* fashion in some implausible theory of volitions. If, however, we view psychological terms as theoretical terms whose employment is neither arbitrary nor *ad hoc*, then there need be no return to an implausible theory of volitions. The use of psychological terms, viewed as theoretical terms in an explanation, helps to fill out the explanation by rendering it more complete, and by providing for systematic interconnections between the observable data and a psychological theory of behaviour.

The reasons in support of this view are analogous to those given for any science where it is maintained that a theory cannot be constructed entirely in terms of observables. Whatever account we adopt concerning the place of theoretical terms in the natural sciences, the same account would seem to be applicable to theories of action and human behaviour. Accordingly, if so-called 'mentalistic' terms can be viewed as theoretical terms, then at least some of the laws governing action will contain such

¹ See, for example, A. I. Melden (1961), especially Chapter V; and G. N. A. Vesey (1961).

terms. So whatever reasons are given (pragmatic or other) for the ineliminability of theoretical terms in general will apply as well to these 'mentalistic' terms which enter into laws governing human action.

The claim that such terms are ineliminable amounts to the view that they are not eliminable in favour of observables or publicly observable behaviour. But this does not imply irreducibility to unobservables or to micro-phenomena such as brain processes or other microphysical states or events. The question remains, however, whether a materialist reduction in terms of microphysical phenomena can be effected—a result which would aid in the provision of laws and in the development of a systematic theory of human behaviour. The answer to this question depends partly on the progress of physiological psychology and related disciplines, and partly on the extra-evidential considerations that need to be invoked in deciding for or against the neural identity theory or some other mind-body theory. In any case, it is far from clear that such things as wants, motives, beliefs, the having of intentions etc., cannot count as causes of action because they do not denote introspectible inner processes of some sort (such as feelings of anger, fear, and the like). It is not necessary to accord the ontological status of an introspectible feeling or inner emotional event to the referents of all psychological terms in order to render them causally relevant factors in explanation of action.

Case Western Reserve University

REFERENCES

- ANSCOMBE, G. E. M. (1963). *Intention*. Ithaca, N.Y.: Cornell University Press.
 BRODBECK, M. (1963). Meaning and action. *Phil. Sci.* 30, 309-24.
 DAVIDSON, D. (1963). Actions, reasons and causes. *J. Phil.* 60, 693-700.
 GEAN, W. D. (1965). Reasons and causes. *Rev. Metaphys.* 19, 667-88.
 GOROVITZ, S. (1965). Causal judgments and causal explanations. *J. Phil.* 62, 695-712.
 HANCOCK, R. (1962). Interpersonal and physical causation. *Phil. Rev.* 71, 369-76.
 MELDEN, A. I. (1961). *Free Action*. London: Routledge & Kegan Paul.
 NAGEL, E. (1961). *The Structure of Science*. New York: Harcourt, Brace & World.
 PERRY, D. L. (1965). Prediction, explanation, and freedom. *Monist.* 49, 234-47.
 SMART, J. J. C. (1964). Causality and human behaviour. *Arist. Soc. Proc. Suppl. Vol.* 38, 143-8.
 TAYLOR, C. (1964). *The Explanation of Behaviour*. London: Routledge & Kegan Paul.
 TAYLOR, R. (1966). *Action and Purpose*. Englewood Cliffs, N.J.: Prentice-Hall.
 VESEY, G. N. A. (1961). Volition. *Philosophy* 36, 352-65.
 YOLTON, J. W. (1966). Agent causality. *Am. Phil. Q.* 3, 14-26.
 YOLTON, J. W. (1966a). My hand goes out to you. *Philosophy* 41, 140-52.

The 'Present' in Physics

by H. A. C. DOBBS

In a recent book entitled *Modern Science and Zeno's Paradoxes*¹ Professor Adolf Grünbaum has put forward the view that the transitory aspect of temporal facts is wholly mind-dependent. This transitory aspect is the one in virtue of which it is usually held by commonsense that physical events do come to pass, or have a transient now-status, in an objective non-mental sense, which they lose after a while when they become wholly past. Grünbaum mentions as examples of such events '(1) a stellar explosion that occurred several million years before time t but which is first seen on earth at time t , and (2) a lightning flash originating only a fraction of a second before t and observed at time t' . He puts his contrary view in the following way :

What qualifies a physical event at a time t as belonging to the present or as now is not some physical attribute of the event or some relation it sustains to other purely physical events; instead what so qualifies the event is that at least one human or other *mind-possessing* organism M experiences the event at the time t , such that at t M is *conceptually aware* of the following complex fact: that his having the experience of the event *coincides temporally* with an awareness of the fact that he has it at all (*loc. cit.* p. 17).

By this assertion Grünbaum seems to be saying that presentness can characterise only those physical events of which at least one 'mind-possessing' organism has some form of awareness both (a) of the event itself and (b) of the awareness of (a); and further the two occurrences (a) and (b) must be contemporary. In fact he goes even further when he asserts that there must be some element of explicit judgment about the awareness by saying :

For example if I just hear a noise at time t , the noise does not qualify at t as *now* unless I am judgmentally aware of the fact of my hearing it at all and of the temporal coincidence of the hearing with that awareness (*ibid.* p. 17).

This view is certainly highly paradoxical. It runs counter to commonsense in the most obvious way. We ordinarily in our practical life act on the belief that we are surrounded by a mass of physical events which are objectively present, quite apart from our perceiving them and making explicit judgments about their happening. It is difficult to see how such

Received 23 July 1968

¹ References are to the U.S. edition, published by Wesleyan University Press. A revised British edition is published by Allen and Unwin.

unpleasant experiences as wholly unforeseen accidents could occur, if only those events of which we were aware were actually happening. However, perhaps the view is worth examining because, despite its apparent silliness (in Broad's sense of 'silly theories' which we could not possibly hold when not philosophising) it has attracted some supporters amongst philosophers, such as Smart, and theoretical physicists, such as Weyl and Eddington.

There are a number of difficulties about Grünbaum's way of asserting the theory, for example his entanglement of the description of the characteristic of 'nowness' or 'presentness' with the very obscure concept of '*mind*-possessing organism'. We do not need to be reminded here of all the complexities involved in the concept of mind. I shall brush these difficulties aside, and assume that a sufficiently unambiguous meaning could be given to Grünbaum's description of his usage, so as to enable the discussion to be kept within reasonable limits, and on a *physical* plane. For I believe his account betrays a serious ignorance of certain disciplines of physics, namely just those in which cognizance has to be taken of 'nowness' (in the sense associated with becoming). He remarks (*loc. cit.* p. 20):

It seems to me of decisive significance that no cognizance is taken of nowness (in the sense associated with becoming) in any of the extant theories of physics. If nowness were a fundamental property of physical events themselves, then it would be very strange indeed that it could go unrecognised in all extant physical theories *without detriment to their explanatory success*.

But there is a very large and important body of physical theory which deals with *prediction* in which the concept of objective physical 'nowness' is basic. I shall content myself with quoting a few illustrative remarks from the work of Norbert Wiener, who did so much to put this branch of physics on a sound footing, to show this. The quotations are from his book *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications* (Wiley: New York, 1949). Throughout this work the 'now' time-point $t=0$ separating the Past (negative values of t) from the Future (positive values of t) is of fundamental significance. This is not merely a matter of diagrammatic convenience (as in the case of Minkowski diagrams), but of the physical applicability of mathematical expressions—specifically those containing symbols called 'operators'—and of their realisability in terms of actual hardware. For example, he remarks (*loc. cit.* p. 8):

The use of the complex plane is intimately allied to the fact that physically applicable operators of engineering allow us to work with the past of our data but not with their future.

And again (*loc. cit.* p. 12):

Not every operator invariant under the translation group is allowed. We have already said that while the past of a time series is accessible for examination, its future is not. That means that our operators must have an inherent one-sidedness. In the case of linear operators an adequate theory of these operators is in existence . . . [i.e. of] all linear operators depending only on the past and invariant under translation [in time].

Wiener points out that much of the significance of the linear translation group, in the so-called 'ergodic' theory of statistical mechanics, lies in its justification for the powerful method of substituting phase averages over a whole contemporary ensemble for the time averages over the behaviour of a single function. He points out (*loc. cit.* p. 47) that

This allows us to identify averages made in the observable past of a time series with averages to be subsequently obtained from the *now* unattainable future.

I think it is obvious from these quotations that contrary to Grünbaum's assertion, there is cognizance of 'nowness' in an important part of physics as it is currently practised by *scientists*. Moreover the theory of prediction, with which Norbert Wiener was concerned in the passages I have quoted, is an important practical part of the physics of communication systems, as well as of meteorology and tidal theory. This consideration must weigh against Grünbaum's approach: since he says that it seems to him 'of *decisive* significance' that 'no cognizance is taken of nowness (in the sense associated with becoming) in any of the extant theories of physics'.

Furthermore the sense of 'presentness' or 'nowness' which is relevant to the physics of prediction is just the very one involving transiency between future and past: in the precise sense that the relevant operators have to be invariant under translation in time but can only operate on the past and not on the future. It is therefore quite a different kind of 'nowness' to the one involved in certain diagrams used to depict spatio-temporal relations along world-lines in Minkowski's work in the special theory of relativity. I agree with Grünbaum that

Minkowski's Here-Now denotes an arbitrarily chosen event of reference, which can be chosen *once and for all* and continues to qualify as now at various times independently of when the diagram is used. Hence there is no transiency of the now in the relativistic scheme depicted by Minkowski; and his absolute past and absolute future are simply absolutely earlier and absolutely later than the arbitrarily chosen fixed reference event called Here-Now (*loc. cit.* p. 15).

The other feature in Grünbaum's account which seems to be true is that, in physics, no cognizance is taken of transient 'nowness' in the sense to which Grünbaum wishes to restrict it: viz. *mental* awareness of a certain

sort. This is not surprising, because physics is not concerned with mental states as such. It is concerned with objective characteristics of the world, which may indeed be the physical counterparts of mental states, but which occur irrespective of whether these are perceived or not. Indeed the 'nowness' which can be said to lie at the heart of the conceptual structure of the physics of statistical predictive systems is a sort of objective counterpart to the human sense of the present; in a way that is analogous to the sense in which the *physical* colour of an object is an objective counterpart to the subjective experience of seeing it as coloured in a certain way.

In the case of colour we have, on the physical side, the so-called 'spectral luminance factor' of an object, as the objective counterpart to the sensible colour it manifests to a particular person—the sensible colour being what Grünbaum would call, I think, a 'mind-dependent' quality. In the case of transiency in time we have the objective $t=0$ division between past and future, in relation to the permissible time-operators of the predictor system, as the objective counterpart to the subjective sense of the passage of time to a particular person. Neither the spectral luminosity of coloured surfaces, nor the transient zero of physical time, is a directly perceptible quantity. Both require specialised physical apparatus for their detection. In the former case a spectrometer to measure spectral densities is required. In the latter case, where for example a predictor system works as part of statistical forecasting equipment, using physical methods dependent as Wiener puts it 'only on the past and present of the function [e.g. the path of the cyclone]', special filters are required. These behave as auto-correlators which are able to yield a prediction by integrating all the data over a given segment of the negative time values of the past up to the time point $t=0$ which is the physical present. The physical 'now', which is the objective counterpart to the 'mental now' of which Grünbaum speaks, consists for such physical predictor systems in the latest time at which information is taken into the leading filter of the predictor.

There is a close analogy to this process in the case of human perception by the central nervous system, which can be regarded as working upon the input, supplied by the peripheral nervous system, with a set of filters. These take, so to speak, a moving time average over a segment of data in the very recent past, integrate it up to the present; and then make a very short 'prediction' ahead of the actual latest instant of the physical 'now': to yield the typical experience of the 'mental present'. It has long been recognised, by writers on the psychology of time, that the so-called specious or mental present is not instantaneous, but comprises both data from the immediate past and also anticipations of the immediate future. For example Professor James Gibson writes:

I would add that we cannot be sure when perception stops and *expectation* begins, either, for a recent experiment at Cornell with an expanding silhouette in the field of view strongly suggests that the perception of approach and the expectation of collision are not separate (*Journal of Psychology*, 62, (1966) 143 referring to W. Schiff (1965), *Psychol. Monog.* 79 no. 604).

So we can fairly say that the position taken by Grünbaum should be reversed. For we can argue, from the admitted neurophysical basis of perception, that the mental sense of transiency of the now must be attributed to an *objective transiency* in the physical processes in the brain of the observer or perceiver. For these processes can be treated as analogous to the kind of processes of averaging over time, and prediction on the basis of a limited section of past data integrated up to the present physical now, found in physical predictor systems. The only way of denying the attribution of objective transient presentness to the physical processes in the brain would be to adopt a kind of extreme mentalism or spiritualism, which denied effective causal relationship between brain processes and mental processes. Such a position is evidently not the one Grünbaum means to adopt *explicitly* (although it might seem implicit in his view) as the following quotation indicates (*loc. cit.* p. 26):

In asserting the mind-dependence of becoming, I allow fully that the kind of neurophysiological brain state which underlies our mere awareness of an event as simply occurring now differs in specifiable ways from the ones underlying tick-tock or melody awareness, memory awareness, anticipation awareness and dream free sleep. But I cannot see why the states of awareness which make for becoming must have physical event counterparts which isomorphically become in their own right.

It is quite clear from this passage that Grünbaum is asserting that all states of mental awareness do have underlying neurophysiological, i.e. *physical* states. This is also confirmed by the footnote on page 17 where he refers to a physical substratum which is, according to all reliable scientific evidence, a necessary condition for consciousness. As to the question why then on this view one must suppose that there are physical event counterparts which 'become', or have a transitory aspect, the answer is that this is the simplest and most obvious explanation of how and why we all have an agreed and working practical simultaneity in our experiences. That there is this practical intersubjective agreement, leading to common acceptance of radio announcers' statements like 'It is now six o'clock in the morning of Monday, 1st January 1968', is perfectly obvious. But unless this agreement is founded on some common objective counterpart to the mental 'now', which *causes* our individual brains to register 'now' with practical simultaneity, it is impossible to see why *I* should not

be contrary, and hold that 'now it is twelve noon on Queen Anne's birthday in the year 1700'. In other words, unless the *mental becoming* of the experience of the specious present is occasioned by the *physical becoming* present of the neurophysical events underlying consciousness, the fact of intersubjective practical simultaneity would be explicable only in terms of a miraculous pre-established (spiritualist) harmony.

The denial of precise distant simultaneity by relativity theory in no way invalidates the foregoing argument. The relativistic modification, of commonsense practical simultaneity, only becomes significant when an attempt is made to establish a common local time point exactly between two observers, moving relatively to each other with a velocity comparable to that of light, i.e. at 186,000 miles a second. This never arises in the case of two human observers on, or near to the surface of this planet, under present conditions, where the maximum attainable relative velocities even by astronauts are measurable in terms of a few thousand miles *per hour*. It is very doubtful if human life could be organised in communities at all if practical simultaneity between observers' frames of reference were not attainable. It is also doubtful if the human organism, whose limit of temporal resolution is of the order of milliseconds, will ever find that the lack of exact simultaneity required by relativity is important in attaining the synchronisation sufficient for practical purposes of agreed experience.

The significance of the foregoing considerations is this. Once it is conceded that neurophysical events, underlying and occasioning the states of awareness which Grünbaum accepts as characterising the 'now', *are* allowed to have transitory characteristics of becoming, then it will be inconsistent to deny this characteristic to a much wider class of physical events. Indeed if we regard the universe as essentially a closed causal system after the manner of Mach, then it would follow that *every* physical event without exception would have the characteristic of presentness at some time or other. For since relations of 'earlier' or 'later' would hold throughout the domain of events, if all events are temporally interrelated by causal chains, it must be true of any event that it is either 'earlier' 'later' or 'compresent' with any event which is now physically present. (I use the word 'compresent' rather than 'strictly simultaneous' to take care of the relativistic restrictions on the time relations with events lying outside the light-cone with respect to a given event.)

Even if we do not accept a closed causal system there will be a vast assemblage of events which have causal relations with events qualifying as having physical 'presentness' because they are physical events which are earlier or later, or compresent with events underlying mental events having the characteristic of subjective experienced presentness.

Perhaps something should be said about what Grünbaum describes as an 'important perplexity which besets the opposing claim that physical events are inherently past present and future'. This perplexity was stated by Smart (*Philosophy and Scientific Realism* p. 135) as follows:

If past, present and future were real properties of events then it would require explanation that an event, which becomes present in 1965, becomes present at that date and not some other (and this would have to be an explanation over and above the explanation of why an event of this sort occurred in 1965).

Unless one is already committed to, and steeped in, the subjectivist notion of becoming present favoured by Grünbaum (and also it would seem by Smart) it is difficult to see the difficulty here. It seems to come from an excessive brooding over the abstract notion of an event *E* in general devoid of any particular content, and thus wondering why *E* should happen now in say 1965 rather than 1700. But if one talks of a particular event, such as the seventeenth birthday of the Prince of Wales who is the son of Queen Elizabeth II, it is obvious that it could only happen in 1965 and not in any other year. Of course, it might be argued that the Prince might have died in 1964, and then his seventeenth birthday would never have become present. But the 'explanation' of why this did not happen, if one is to be required, seems to belong to the medical history of Prince Charles, and to have nothing whatever to do with the question of whether 'past, present and future' are real properties of events, or are only mind-dependent characteristics ascribed to events by particular persons.

Finally, I think one should attempt to understand why so odd a theory as the denial of the objectivity of the happening of physical events should be put forward. As I said at the outset, it seems to me to be an instance of what C. D. Broad called a 'silly' theory, in that it could not possibly be held practically in the sense of being acted upon when one is not actually philosophising. Broad made it clear that a 'silly' theory in this sense might be both ingenious, and held in the abstract by highly intelligent people. In fact the denial of the physical objectivity of becoming or happening of events has held a strange fascination for some of the most profound intellects of our time. I am thinking in particular of Weyl and Eddington, perhaps after Einstein himself, the two outstanding exponents of general relativity theory in its initial phase. Eddington, however, had a strong and acknowledged vein of mysticism, which comes out explicitly in his famous Swarthmore Lecture entitled *Science and the Unseen World*. Weyl wrote in terms very similar to Eddington, as may be seen in the following quotation, from his Lectures delivered under the auspices of the

William J. Cooper Foundation of Swarthmore College, published under the title *Mind and Nature* (Philadelphia, 1934), p. 76:

The objective world merely exists, it does not happen; as a whole it has no history. Only before the eye of consciousness climbing up in the world line of my body, a section of this world comes to life and moves past it as a spatial image engaged in a temporal transformation.

While philosophers may be forgiven intellectual extravagances of this kind, I think it is a pity when they receive encouragement from theoretical physicists, even if these share a common mystical experience.

Travellers Club, London

The Structure of Acceptance and its Evidential Basis

by P. M. WILLIAMS

The object of this study is first to analyse the logical structure of rational acceptance. Axiom systems representing strong and weak concepts will be distinguished, but will subsequently be shown essentially equivalent with respect to the concepts of positive acceptance, positive rejection, and suspension of judgment, to be defined later. Secondly, in an application of these ideas, a pragmatistical analysis of the concept of evidence will be provided using, as far as possible, the proposed structure of acceptance alone.¹

I INTRODUCTION

Acceptance is considered in this paper as a classificatory concept: no mention will be made of *degrees* of acceptance. The importance of such a concept for the methodology of science has been the subject of some debate (see, e.g., Rudner, 1953; Jeffrey, 1956; Levi, 1960; Carnap, 1964, pp. 972-3; Levi, 1967, pp. 7-21). A solution to the central problem of this debate requires a precise definition of the concept of acceptance under discussion and a clear demarcation of the process of science. I shall attempt neither of these here. Instead, I shall merely suppose, without discussion, that scientific hypotheses can meaningfully be said to be accepted in some sense and, therefore, that a classificatory concept is itself at least meaningful. The question as to whether it is of importance for the methodology of science will be left open. But I hope to show later that it can be of interest for an analysis of the concept of *evidence*. It may well be of wider epistemological interest.

Although scientific hypotheses are often said to be 'generally accepted' at some time, this is usually understood to be an elliptical way of saying that they are accepted at that time by a substantial majority of those thought competent to judge them. In this expanded form the reference to persons is evident. A hypothesis cannot just be accepted: there has to be someone to do the accepting and this must take place at some definite

Received 23 July 1968

¹ I am grateful for many valuable discussions with Dr Jerzy Giedymin during the preparation of this work and also to Dr Aaron Sloman for commenting on an earlier draft.

moment or period in time. Being a relation between the sentences of some language and one or more users of that language, acceptance may be termed a *pragmatical* relation. .

Although no ultimate commitment will be made to a unique interpretation or rule of acceptance, it will be necessary when presently considering the intuitive adequacy of possible axioms to have some such rule in mind. For this purpose I shall suppose that the agent X accepts at time t all and only those sentences of his language in which he has a sufficiently high degree of belief at time t .

Strong arguments have been put forward for requiring that the measure P , representing X 's degrees of belief at any one time, should be a normalised probability measure if his simultaneously held beliefs are to be *coherent*, or *rational* (see, e.g., de Finetti, 1937; Carnap, 1962, pp. 307-8). Since this study attempts to outline a concept of rational acceptance, it is reasonable to demand that P satisfy the following:

(P1) $P(h)$ is a non-negative real number,

(P2) $\vdash (\sim(h \& i)) \rightarrow P(h \vee i) = P(h) + P(i)$,

(P3) $\vdash (h) \rightarrow P(h) = 1$.

(The symbol ' $\vdash$ ' refers to the formula immediately succeeding it in parentheses.) The following consequences of (P1-3) will be needed:

$$P(h) + P(\sim h) = 1, \quad (1)$$

$$P(h) \leq 1, \quad (2)$$

$$\vdash (h \leftrightarrow i) \rightarrow P(h) = P(i), \quad (3)$$

$$\text{and} \quad P(h \vee i) = P(h) + P(i) - P(h \& i). \quad (4)$$

P is the probability measure appropriate to X at time t . (Subscript references to X and t have been omitted for notational simplicity.) Nevertheless, although we have spoken above of 'degrees of belief', the interpretation of probability associated with P need not be subjective. $P(h)$ may be identified, definitionally, with $P'(h, e)$ where P' is any probability measure thought appropriate, interpreted in any of the customary ways, logical, frequency, or otherwise, and e is the (fixed) total evidence available to X at time t . Provided only that P' satisfies the usual axioms for conditional probability, P will satisfy (P1-3).

The working concept of acceptance suggested above may now be made more precise in the following probabilistic rule:

(PRA) $A(h) \leftrightarrow P(h) \geq 1 - \epsilon$, where $0 \leq \epsilon \leq 1$.

When considering the adequacy of possible axioms I shall adopt the criterion (C) that an axiom for acceptance is satisfactory if and only if it is consistent with (PRA) for some specified value, or range of values, of ϵ in $[0, 1]$.

It is important to remember that when axioms have been selected with the help of this criterion, they may be assessed independently of it; for the predicate 'A' may be interpreted in ways other than that required by (PRA). Whereas the axioms we shall admit, using (C) as a heuristic guide, seem to be fairly uncontroversial even for those preferring different interpretations of acceptance, the axioms we shall disallow include some that may appear equally satisfactory. (C) may therefore be thought too restrictive in sanctioning only an (intuitively) incomplete set of axioms. Nevertheless, even this set appears to have interesting properties.

2 LANGUAGE AND META-LANGUAGE

In his book *Toward a Systematic Pragmatics*, Martin (1959) has formulated a pragmatical meta-language in which to include a relation of acceptance. It is assumed that the formal part of this study could be incorporated into a system such as Martin's, a few of whose characteristics will now be sketched together with certain, mainly notational, departures that will be made from it.

Without actually specifying it, I shall assume that the logical structure of the object-language L is given. It is presumed to include logical rules providing for the theory of truth-functions.

The syntax of L may be described in the syntactical meta-language, which is constructed for that purpose. It contains names of all the individual sentences of L . 'a', 'b', 'c', etc., with or without numerical subscripts, are taken as sentential variables ranging over these names. For simplicity the conventions are adopted here that symbols of the object-language may also be used as names for themselves, and that compound names are formed in the usual way by juxtaposition.

The syntactical meta-language is incorporated in the semantical meta-language, which in turn is incorporated in the pragmatical meta-language. The rules of the meta-languages include logical rules providing for the theory of truth-functions, and appropriate rules of quantification.

Whereas Martin writes, in the pragmatical meta-language, ' X Acpt h, t ' to mean 'the human being X accepts the sentence h at time t ', I shall simply write ' $A(h)$ ', making explicit reference to the person X and the time t only when necessary.

The following *Rule of Restriction* is introduced by Martin (1959, p. 37):

$$X \text{ Acpt } h, t \rightarrow \text{Sent } h.$$

This is presupposed here: only sentences of L are to be accepted.¹ ('S_L')

¹ The requirement that only sentences (and not, for example, propositions) are to be accepted means that an *extensional* meta-language may be used (q.v. Martin, 1959, pp. 5-6).

will stand for the class of all sentences of L , to which informal reference will sometimes be made.)

Martin introduces into the pragmatical meta-language certain abbreviations. The following are paraphrases¹ of the definitions of two such abbreviations (p. 40):

X will be said to accept ' $\sim$ ' in its normal sense at time t if and only if, for every sentence h , X accepts h at t if and only if he does not accept $\sim h$ at t ; (5)

and

X will be said to accept ' $\vee$ ' in its normal sense at time t if and only if, for every sentence h , and for every sentence i , X accepts $h\vee i$ at t if and only if he accepts h or he accepts i at t . (6)

Martin uses these for the purpose of classifying the *users* of the object-language. It is therefore an empirical question as to whether any given person X does or does not "accept ' $\sim$ ' in its normal sense at time t ", etc. Since this study attempts to outline a concept of *rational* acceptance, it is more natural to lay down postulates concerning X 's acceptances and non-acceptances at time t . It is then an empirical question as to whether the calculus to be suggested here represents, as a whole, the behaviour of a given person at a given time. It would be possible to reformulate all that follows in the descriptive approach of Martin but a normative approach is simpler.

(5) and (6) suggest that the following two conditions should be satisfied by a concept of rational acceptance:

$$A(h) \leftrightarrow \sim A(\sim h), \quad (7)$$

and

$$A(h\vee i) \leftrightarrow A(h)\vee A(i). \quad (8)$$

(The universal quantifiers that should occur at the beginnings of formulae such as (7) and (8) are omitted for brevity. Every formula containing two or more occurrences of the same apparently free sentential variable is understood to be within the scope of a universal quantifier binding that variable. It should also be noted that, in accordance with our conventions on naming, the symbol ' $\vee$ ' appears in (8) in two capacities. On the right of the sign for the biconditional it occurs as the disjunctive connective of the meta-language. On the left it appears as the *name* of the corresponding connective of the object-language. The context should keep this distinction clear. Analogous remarks hold for the other connectives.)

Let us suppose for the moment that (7) and (8) are introduced as axioms,

¹ Martin relativises acceptance to any finite (or denumerably infinite) class of sentences under investigation relative to one or more persons (p. 38). Mention of this class is made in these abbreviations in their original form, but it may be omitted here without doing violence to their original meanings.

and also that X does not distinguish, in relation to acceptance, between logically equivalent sentences. This may be expressed by

$$\vdash (h \leftrightarrow i) \rightarrow (A(h) \leftrightarrow A(i)).$$

(7) and (8), together with this condition, yield

$$A(h \& i) \leftrightarrow A(h) \& A(i), \quad (9)$$

and

$$A(h \rightarrow i) \leftrightarrow (A(h) \rightarrow A(i)). \quad (10)$$

Martin's use of the locution 'in its normal sense' in (5) and (6) suggests that anyone not satisfying (7) and (8), and by implication (9) and (10), must be using sentential connectives in a non-standard way. It would appear, therefore, that the acceptances or non-acceptances of a rational person ought to satisfy these conditions. The contention here is that it is possible for a person to be rational to the extent that his degrees of belief satisfy (P1-3) (which indicate that his use of sentential connectives is far from being non-standard) and yet for him to fail to satisfy all of (7)-(10). Furthermore, his rationality, in this sense, may even preclude his satisfaction of the complete set of conditions if his acceptances are related to his degrees of belief by (PRA).¹ The next section will therefore contain a comprehensive survey of the conditional components of the biconditionals (7)-(10) to discover under what conditions, if any, they are compatible with (PRA).

3 THE STRUCTURE OF ACCEPTANCE

We have first

$$A(h) \rightarrow \sim A(\sim h). \quad (7a)$$

This makes the apparently reasonable demand that X at no time accept both members of a pair of contradictory statements. If $\epsilon < \frac{1}{2}$, (7a) can be proved with the help of (PRA). For if the antecedent of (7a) holds, the probability of h must be no less than $1 - \epsilon$. Thus by (1) the probability of $\sim h$ must be no greater than ϵ which is strictly less than $1 - \epsilon$ if $\epsilon < \frac{1}{2}$. Thus $\sim h$ cannot be accepted and the proof is complete. If, on the other hand, $\epsilon \geq \frac{1}{2}$, (7a) can be shown to be incompatible with (PRA). For it is not unreasonable to suppose that there is some sentence f of L satisfying

$$\epsilon \geq P(f) \geq 1 - \epsilon. \quad 2$$

¹ This means that Martin's concept of acceptance is not conformable with (PRA). Nevertheless, with respect to one of his own suggested interpretations of acceptance (p. 10), it would appear that even those with an impeccable knowledge and understanding of sentential logic may fail to be classified as accepting the sign for negation, at least, in its 'normal sense'. For this to happen they need only fail to answer 'Yes' both to the question 'Does a hold?' and to the question 'Does $\sim a$ hold?' for some sentence a ; that is to say, they need only suspend judgment about some sentence.

² I shall make several suppositions of this sort concerning the availability in L of sentences with probability values lying in specified ranges. For almost every probability distribution, languages in which these are not satisfied will be wholly inadequate for the purposes of science.

Then, from (1),

$$P(\sim f) \geq 1 - \epsilon,$$

and so by (PRA) both f and $\sim f$ are accepted. This contradicts (7a). Therefore, in conformity with the criterion (C), (7a) is satisfactory if and only if $\epsilon < \frac{1}{2}$.

For the next condition,

$$\sim A(\sim h) \rightarrow A(h), \quad (7b)$$

the exact converse holds. (7b) is satisfactory if and only if $\epsilon \geq \frac{1}{2}$. This can be shown in a similar way.

The next condition is

$$A(h \vee i) \rightarrow A(h) \vee A(i).^1 \quad (8a)$$

This seems too strong. For consider the two sentences 'The next toss of this coin will yield heads' and 'The next toss of this coin will yield tails'. Then, in most circumstances, it would surely not be unreasonable to accept the disjunction of these two sentences and yet at the same time to accept neither sentence individually. This would conflict with (8a). In probabilistic terms, if $\epsilon < 1$, it is consistent with (P1-3) to suppose that there are two sentences of L , each having a probability strictly less than $1 - \epsilon$, yet whose disjunction is at least as probable as $1 - \epsilon$. Given (PRA) this contradicts (8a). (8a) is shown by this argument to be unsatisfactory for all values of ϵ other than 1. The case in which $\epsilon = 1$ will be considered shortly.

The next two conditions,

$$A(h) \vee A(i) \rightarrow A(h \vee i) \quad (8b)$$

and

$$A(h \& i) \rightarrow A(h) \& A(i), \quad (9a)$$

may be taken together. They both seem intuitively reasonable. For, in the case of (8b), if X accepts h or if he accepts i , he may reasonably be expected to accept the weaker sentence $h \vee i$ which is logically entailed by both. For (9a), h and i are each logically entailed by $h \& i$, and therefore both the acceptance of h and the acceptance of i may be expected to follow from the acceptance of the stronger sentence $h \& i$. In probabilistic terms, if p logically entails q , from (P3)

$$P(\sim p \vee q) = 1. \quad (11)$$

But from (4)

$$P(\sim p \vee q) = P(\sim p) + P(q) - P(\sim p \& q). \quad (12)$$

Therefore from (11), (12) and (1) we have

$$P(p) = P(q) - P(\sim p \& q);$$

so, from (P1)

$$P(p) \leq P(q).$$

Given (PRA), this justifies, independently of the value of ϵ ,

$$\vdash (p \rightarrow q) \rightarrow (A(p) \rightarrow A(q)), \quad (13)$$

¹ A similar principle has been proposed by Chisholm (1957, p. 13) and criticised by Seigh (1964).

from which both (8b) and (9a) follow. In accordance with (C) therefore, (13), and hence (8b) and (9a), are satisfactory for all values of ϵ in the closed interval $[0, 1]$.

Since each sentence of L logically entails every logical truth of L , it follows from (13), and the assumption that X accepts at least one sentence of L , that X accepts all logical truths of L :

$$(\exists e)(A(e) \rightarrow (h)(\vdash(h) \rightarrow A(h))). \quad (14)$$

Next we have

$$A(h) \& A(i) \rightarrow A(h \& i). \quad (9b)$$

This is an important condition. For to postulate this and (13), which has been found satisfactory for all values of ϵ , would ensure that the class of sentences accepted by X at any one time possessed the attractive property of deductive closure. Nevertheless, as is well known, (9b) is incompatible with (PRA) for all values of ϵ in the open interval $(0, 1)$. For it is consistent with the axioms of probability for each of two sentences to have probabilities greater than or equal to $1 - \epsilon$ and yet for the probability of their conjunction to be strictly less than $1 - \epsilon$. Given (PRA) this contradicts (9b). Thus (9b) is unsatisfactory for all values of ϵ in $(0, 1)$.¹

On the other hand, (9b) is consistent with (PRA), and indeed deducible from it given the axioms of probability, if ϵ takes either of the end-point values, 0 or 1. If X assigns probability 1 to all and only those sentences of

¹ It is well known that any purely probabilistic concept of acceptance satisfying (PRA) is incompatible with the first two of three 'conditions of rationality in the formation of beliefs' that have been proposed by Hempel (1962, pp. 150-1). These two conditions require the class of accepted statements to be deductively closed and logically consistent, respectively. That (PRA) is incompatible with (CR1) is shown by the discussion of (9b) above. The incompatibility of (PRA) with (CR2) is shown by a similar argument (q.v. Hintikka & Hilpinen, 1966, p. 5). Nevertheless, (PRA) is compatible with Hempel's conditions if these are reinterpreted in the manner suggested by Kyburg (1965, p. 307)—and this by virtue of (13) and (7a) respectively, which have been shown to hold for appropriate values of ϵ . Several writers, notably Hintikka & Hilpinen (1966) and Levi (1967), however, have been led to a consideration of rules of acceptance other than (PRA) by the wish to satisfy (CR1) and (CR2) when these are interpreted in the strong sense in which they were originally intended. This has meant the abandonment of purely probabilistic rules of acceptance. Hintikka & Hilpinen (1966) have provided a definition of acceptability (p. 11) which takes account not only of the probability of the hypothesis in question, on the basis of the available evidence, but also of the number of individuals referred to in the evidence, i.e. of the size of the observed sample. (Probability is measured here with the help of Hintikka's system of inductive logic; but for this to be applicable it is necessary to make the strong assumption that the sentence expressing the available evidence will, in effect, say for each individual mentioned in it and for every property denoted in the object-language, whether or not that individual possesses that property: q.v. p. 7.) Levi (1967), on the other hand, makes use of a rule of rejection corresponding to (PRA) but one in which the level of rejection is not a preassigned constant but rather a variable depending on the content of the hypothesis concerned. The resulting inductive acceptance rule (p. 86), relativised to the set of 'relevant answers' to a given problem, is then compatible with (CR1) and (CR2) if their scope is restricted, in each case, to the set of sentences accepted relative to the problem in question (q.v. pp. 32-6, 91-5).

L which he unreservedly believes to be true, the case $\epsilon = 0$ may be interpreted as corresponding to the concept of acceptance *as true*.¹ On some views of acceptance it might be argued that the only sentences which may legitimately, or rationally, be accepted are those that can be accepted as true (in this sense). The class of sentences thus accepted, however, will be extremely limited. It will contain the truths of logic and mathematics but little else besides. (It might contain 'observation' or 'evidence' sentences.) Since I do not wish to exclude from consideration more liberal concepts of acceptance, except those that are so liberal as to allow every sentence of L to be accepted (e.g. when $\epsilon = 1$), I shall strengthen (C) by requiring that no axiom for acceptance be considered satisfactory, for present purposes, unless it is consistent with (PRA) for some range of values of ϵ in the *open* interval $(0, 1)$. Thus (9b) is unsatisfactory.

The next condition is

$$A(h \rightarrow i) \rightarrow (A(h) \rightarrow A(i)). \quad (10a)$$

This may be seen as a pragmatist correlate of the principle of *Modus Ponens* (cf. Martin, 1959, p. 41). Nevertheless, it is incompatible with (PRA) for all values of ϵ in $(0, 1)$. For it is consistent with the axioms of probability to suppose that there are sentences f and g of L such that

$$P(f) = 1 - \lambda\epsilon, \quad (15)$$

$$P(g) = (1 - \mu)(1 - \epsilon), \quad (16)$$

and
$$P(f \vee g) \leq 1 - \mu(1 - \epsilon), \quad (17)$$

for some real numbers λ and μ satisfying

$$(i) \quad 0 < \lambda \leq 1,$$

and
$$(ii) \quad 0 < \mu \leq \min\{1, \lambda\epsilon/(1 - \epsilon)\}.$$

From (16), (17) and the probability calculus, it follows that

$$P(\sim f \vee g) \geq 1 - \epsilon. \quad (18)$$

Therefore, since $\lambda \leq 1$ and $\mu > 0$, using (15), (16), (18), (3), and (PRA), we have

$$(\exists f)(\exists g)(A(f \rightarrow g) \& A(f) \& \sim A(g)),$$

which contradicts (10a). (10a) is therefore unsatisfactory.²

The last of the eight conditions,

$$(A(h) \rightarrow A(i)) \rightarrow A(h \rightarrow i), \quad (10b)$$

¹ This concept appears to be nearer the one Martin (1959) has in mind when he writes (p. 11): 'To say that X accepts a sentence a means essentially that . . . he, X , consciously takes a to be true or regards a as true when a is presented to him in an appropriate way.'

² (10a) is, however, satisfied for our concept of acceptance *as true*, not all of (7a)-(10b) are satisfied. The rejection of (10a) for weaker concepts of acceptance is not in conflict with one of the *deductive meaning-rules* proposed by K. Ajdukiewicz (see, e.g., Ajdukiewicz, 1949, p. 182).

is satisfactory if and only if $\epsilon \geq \frac{1}{2}$. It is entailed by (13) and (7b), and is therefore redundant if these are postulated.

This concludes the discussion of (7a)-(10b). There are, however, two interesting conditions that have not yet been mentioned. The first of these is

$$A(h \vee i) \rightarrow \sim A(\sim h) \vee \sim A(\sim i). \quad (19)$$

If $\epsilon < \frac{1}{3}$, (19) is entailed by (PRA) and (P1-3). For suppose that the consequent of (19) does not hold. This is equivalent to supposing that both $\sim h$ and $\sim i$ are accepted. Then from (PRA) and (1)

$$P(h) \leq \epsilon \text{ and } P(i) \leq \epsilon, \quad (20)$$

whence from (4), (P1), (20) we have

$$P(h \vee i) \leq 2\epsilon \\ < 1 - \epsilon, \text{ since } \epsilon < \frac{1}{3}.$$

Thus $h \vee i$ is not accepted and the proof is complete. If $\epsilon \geq \frac{1}{3}$, it can be shown that (19) is incompatible with (PRA). (19) is therefore satisfactory if and only if $\epsilon < \frac{1}{3}$.

The other remaining condition is

$$\sim A(\sim h) \& \sim A(\sim i) \rightarrow A(h \& i). \quad (21)$$

(21) can be shown satisfactory, in a similar way, if and only if $\epsilon \geq \frac{2}{3}$.

As a result of this discussion, the following two axiom systems for acceptance are proposed:

$$\begin{aligned} (\alpha A1) \quad & A(h) \rightarrow \sim A(\sim h), \\ (\alpha A2) \quad & \vdash (h \rightarrow i) \rightarrow (A(h) \rightarrow A(i)), \\ (\alpha A3) \quad & A(h \vee i) \rightarrow \sim A(\sim h) \vee \sim A(\sim i); \end{aligned}$$

and

$$\begin{aligned} (\beta A1) \quad & \sim A'(\sim h) \rightarrow A'(h), \\ (\beta A2) \quad & \vdash (h \rightarrow i) \rightarrow (A'(h) \rightarrow A'(i)), \\ (\beta A3) \quad & \sim A'(\sim h) \& \sim A'(\sim i) \rightarrow A'(h \& i). \end{aligned}$$

A purely probabilistic interpretation of 'A' of the α -axioms is consistent with (P1-3) (i.e. with the coherence of the agent's beliefs) if and only if $0 \leq \epsilon < \frac{1}{3}$. The same type of interpretation of 'A'' of the β -axioms is similarly consistent if and only if $\frac{2}{3} \leq \epsilon \leq 1$. There is no need, of course, to give either 'A' or 'A'' a probabilistic interpretation.

4 SUSPENSION OF JUDGMENT

So far in this study the options open to the agent X have been limited to acceptance or non-acceptance. I now wish to extend these possibilities by supposing, instead, that there are three cognitive relations in which X can stand to the sentences of L . (These relations are taken, for the moment,

to be unrelated to those previously mentioned.) They are *positive acceptance*, *positive rejection*, and *suspension of judgment*, which, although relations between X and the sentences of L at some time t , may be represented by the one-place predicates ' PA ', ' PR ', and ' SJ ' if the implicit references to X and t are understood. For any person and for any sentence of L it is required that one and only one of these relations holds at a given time. Thus the extensions of ' PA ', ' PR ', and ' SJ ' are mutually exclusive and jointly exhaustive over S_L . The question arises concerning the structural conditions to be satisfied by the three new relations. It would be possible to consider, as in the last section, conditions such as (7a)-(10b) with ' A ' replaced by each of the new predicates in turn. This would be lengthy and partly unnecessary, in view of the requirement made above, since any one of these predicates may be defined in terms of the other two; for example

$$SJ(h) \leftrightarrow \sim PA(h) \& \sim PR(h). \quad (22)$$

Thus only two need be taken as primitive. It would be more satisfactory, however, if only one need be taken as primitive, or, equivalently, if all three could be defined with respect to one auxiliary term.

It is suggested that there are two alternative, but essentially equivalent, ways of effecting this reduction: one uses the predicate ' A ' of the α -axioms and the other uses ' A ' of the β -axioms. These are related to ' PA ', ' PR ', and ' SJ ' by the following *alternative* sets of definitions:

$$\begin{aligned} (\alpha D_1) \quad & PA(h) \leftrightarrow A(h), \\ (\alpha D_2) \quad & PR(h) \leftrightarrow A(\sim h), \\ (\alpha D_3) \quad & SJ(h) \leftrightarrow \sim A(h) \& \sim A(\sim h); \end{aligned}$$

or

$$\begin{aligned} (\beta D_1) \quad & PA(h) \leftrightarrow \sim A'(\sim h), \\ (\beta D_2) \quad & PR(h) \leftrightarrow \sim A'(h), \\ (\beta D_3) \quad & SJ(h) \leftrightarrow A'(\sim h) \& A'(h). \end{aligned}$$

If the α -definitions are used, ' A ' is to be governed by the α -axioms, and if the β -definitions are used, ' A ' is to be governed by the β -axioms. The underlying pragmatical meta-language augmented by the α -definitions and the α -axioms will be called the α -system. The β -system is defined correspondingly.

The two systems are claimed to be essentially equivalent representations of positive acceptance, positive rejection, and suspension of judgment, since it is possible to define a purely syntactical one-to-one transformation mapping the (well-formed) formulae of the α -system onto those of the β -system (and vice versa), which furthermore maps theorems on to theorems.

To help define this transformation, each occurrence of the symbol '*A*' in any formula of the α -system will be classified by an ordered pair of natural numbers. An occurrence of '*A*' will be said to be an ' (m, n) -occurrence' if there are (i) just m occurrences of ' $\sim$ ' preceding '*A*' between '*A*' and the last occurrence before it of any symbol other than ' $\sim$ ', or '(' if occurring inessentially, if there is such an occurrence—otherwise, between '*A*' and the beginning of the formula concerned; and (ii) just n occurrences of ' $\sim$ ' between the symbol '(' immediately succeeding '*A*' and the next occurrence after it of any symbol other than ' $\sim$ ', or '(' if occurring inessentially, if there is such an occurrence—otherwise, between '*A*' and the end of the formula concerned. Occurrences of '*A*' in formulae of the β -system are classified similarly.

An example should make this clear. In particular, I hope it will explain what is meant here by an inessential occurrence of '(' without the need for an elaborate, though purely syntactical, definition.¹ Thus the exhibited occurrence of '*A*' in

$$((\dots \&(\sim(\sim A(\sim(\sim h)))))) \rightarrow (\dots))$$

is a $(2, 2)$ -occurrence, whereas that in

$$((\dots \&A(\sim((\sim h)\vee(\sim i)))) \rightarrow (\dots))$$

is a $(0, 1)$ -occurrence.

The transformation d is now defined as follows. For any formula W of the α -system, $d(W)$ is the formula of the β -system which is identical to W except for containing an $(m \pm 1, n \pm 1)$ -occurrence of '*A*' wherever W contains an (m, n) -occurrence of '*A*', the plus and minus signs being taken according as m and n are severally even or odd respectively. (0 is taken as even.)

Since the syntactical rules of formation are the same for the two systems, and since their symbols differ only in respect of '*A*' and '*A*', it is clear that d maps the formulae of the α -system *onto* those of the β -system. It is also clear that d is a one-to-one transformation whose inverse d^{-1} is such that, for any formula W' of the β -system, $d^{-1}(W')$ is the formula of the α -system which is identical to W' except for containing an $(r \pm 1, s \pm 1)$ -occurrence of '*A*' wherever W' contains an (r, s) -occurrence of '*A*', etc. As a trivial example of these one-to-one transformations, it can easily be seen that (αA_1) corresponds to (βA_1) .

The detailed proof of the following meta-theorem is somewhat lengthy. The result will therefore be stated without proof.

¹ The distinction between essential and inessential occurrences of parentheses is, of course, only appropriate to a system of notation of the present type. In the Łukasiewicz system all occurrences of parentheses are inessential and, indeed, in that system our classification could be effected with much less trouble.

For any formula W of the α -system, W is a theorem of the α -system if and only if $d(W)$ is a theorem of the β -system; conversely, for any formula W' of the β -system, W' is a theorem of the β -system if and only if $d^{-1}(W')$ is a theorem of the α -system.

The two systems are therefore *dual* with respect to d .

Bearing in mind that d and d^{-1} , restricted to formulae in which there are no occurrences of ' A ' or ' A' ' respectively, are *identity* transformations, it follows that the theorems involving only ' PA ', ' PR ', and ' $S\mathfrak{f}$ ' are the same in both systems. This means that these predicates may be given the same interpretations in the two systems provided that ' A ' and ' A' ' are either left uninterpreted or, as it turns out, if they are interpreted in such a way as to satisfy (αDI) and $(\beta DI')$

$$\sim PA(\sim h) \leftrightarrow A'(h),$$

respectively. The two systems may therefore be used equivalently to represent the same concepts of positive acceptance, positive rejection, and suspension of judgment.

Before considering the significance of this equivalence, two theorems (of both systems) deserve brief mention. In the first place we have

$$PR(h) \leftrightarrow PA(\sim h). \quad (23)$$

This reflects the possibly unwanted additional structure that has been given to ' PA ' and ' PR ' by defining them in terms of one auxiliary predicate, viz. ' A ' or ' A' '. For suppose that ' PA ' and ' PR ' are to be interpreted purely probabilistically such that, for some specified values of ζ and η ,

$$(PRPA) \quad PA(h) \leftrightarrow P(h) \geq 1 - \zeta$$

$$\text{and} \quad PR(h) \leftrightarrow P(h) \leq \eta.$$

If the error of positively accepting hypotheses when they are false is considered more serious, in some sense, than that of positively rejecting them when they are true, it would be reasonable to want to take ζ to be strictly less than η . (23), however, requires them to be equal, thereby limiting the possible interpretations of ' PA ' and ' PR '. This seems a fair price to pay for the formal simplicity of using only one primitive predicate.

Secondly we have

$$PA(h \vee i) \rightarrow \sim (PR(h) \& PR(i)), \quad (24)$$

i.e. if the disjunction $h \vee i$ is positively accepted, it must not be the case that both h and i are positively rejected. This seems reasonable enough. Corresponding to our discussion of (19), however, it can be seen that whenever a purely probabilistic concept of positive acceptance satisfying (PRPA) is envisaged, (24) will require the level of acceptance to be greater than $\frac{2}{3}$, i.e. ζ must be strictly less than $\frac{1}{3}$.¹

¹ This minimum level of acceptance is higher than is often stipulated; see, e.g., Chisholm (1957, p. 28), Hempel (1962, p. 155), and Hintikka & Hilpinen (1966, p. 11).

Let us call concepts of acceptance satisfying ($\alpha A1-3$), *strong* concepts of acceptance, and those satisfying ($\beta A1-3$), *weak* concepts of acceptance. Now if we intend to give 'PA', 'PR', and 'S' the same interpretations in both systems we may conjoin their axioms and definitions to form the combined $\alpha\beta$ -system. Within this we shall have both

$$A(h) \leftrightarrow \sim A'(\sim h) \quad (25)$$

and

$$A'(h) \leftrightarrow \sim A(\sim h). \quad (26)$$

In view of (25) and (26) we may associate pairs of strong and weak concepts of acceptance in such a way that a hypothesis is accepted in some given strong sense if and only if its negation is not accepted in the associated weak sense. Conversely, a hypothesis is accepted in some given weak sense if and only if its negation is not accepted in the associated strong sense.¹

These relations between associated strong and weak concepts of acceptance are formally analogous to those holding between the *alethic* modal terms 'necessary' and 'possible'. There is a similar analogy with the *epistemic*, *deontic*, and *existential* modalities discussed by von Wright (1951). To pursue this analogy briefly, I shall only discuss formal similarities with the deontic modalities, which are concerned with the concepts of obligation and permission. These modalities are chosen since concepts of acceptance share with them their unrelatedness to matters of fact. There is no logical relation between an act's being obligatory or permitted and its being performed or omitted. Similarly, there is no logical relation, much as one might like there to be, between a sentence's being accepted or rejected and its being true or false. In this respect the deontic modalities differ from each of the other types of modality mentioned above (cf. von Wright, 1951, p. 41).

Prior (1962, p. 221) has restated as follows (using the Łukasiewicz system of notation) certain principles of deontic logic proposed by von Wright (1951, ch. v):

Using 'a', 'b', 'c', etc., for assertions of the form 'An act of the sort A is done', 'P' for 'It is permissible that (such-and-such an act be done)', and 'O' for 'It is obligatory that',² quite a substantial 'deontic logic' may be built up by dovetailing into the propositional calculus the two special axioms

AD1. $COaPa$

AD2. $E(PAab)(APaPb)$,

¹ For a similar distinction between strong and weak senses of *acceptability*, see Hintikka & Hilpinen (1966, p. 2, n. 2).

² In this respect Prior's presentation differs from von Wright's (1951, p. 36) and to a greater extent from that in von Wright (1963), where the latter has subsequently developed a 'Logic of Action' as a prerequisite for a more adequate statement of deontic logic (see especially pp. v-vii).

the definition

Df. $O: O = NPN$,

and the rule

RD1. $E\alpha\beta \rightarrow EP\alpha P\beta$.

If we now let 'O' play the part of 'A', and 'P' that of 'A'', our $\alpha\beta$ -system gives rise to a comparable deontic logic:

AD1'. $COaPa$,

AD2'. $C(OAab)(APaPb)$,

Df. $O: O = NPN$,

RD1'. $C\alpha\beta \rightarrow CP\alpha P\beta$.

RD1' is stronger than RD1, but AD2' is weaker than AD2.

We see now that the α -system corresponds to a deontic logic formulated wholly in terms of obligation, and the β -system to one formulated wholly in terms of permission. The equivalence between the two systems with respect to the transformation d corresponds to the equivalence between these two ways of formulating deontic logic, given the usual relation between obligation and permission.¹

Returning to our intended interpretations of 'A' and 'A'' as strong and weak concepts of acceptance respectively, we see that although these two classes of concepts differ radically in respect of the structural conditions satisfied by their members, they may be used equivalently to define the same classes of concepts of positive acceptance, positive rejection, etc. We are therefore free to choose either system to work with. In the following application of the above analysis to the problem of evidence, I shall choose to use the α -system. No further use will be made of 'PA', 'PR', and 'SY'. Instead, I shall understand 'A(h)' as holding if and only if X (positively) accepts h at time t , 'A($\sim h$)' as holding if and only if X (positively) rejects h at time t , and ' $\sim A(h) \& \sim A(\sim h)$ ' as holding if and only if X suspends judgment about h at time t .

5 EVIDENCE

It is usual to suppose that a certain subclass of the sentences of the language of science may be known, or decided, more or less directly on the basis of experience alone. These sentences are, furthermore, understood to constitute in turn the only appropriate basis with which to confront, and on which thereby to assess, those of the remaining sentences in which we are interested. The process of confrontation between sentences of these two types will not concern us. Rather, our concern is only to draw as precise

¹ I am grateful to the referee for bringing this point to my attention.

a line as possible between them. This we shall attempt by analysing some pragmatical features of sentences of the first type, which will be called 'evidential sentences'.

Such sentences have sometimes been required by philosophers to satisfy strict conditions—that of being decidable by direct observation, for example—in order to ensure that they may be known to be true or false, as the case may be, with *certainty*. Although this understanding of evidential sentences is compatible with the axioms we shall propose, our axioms are intended to require, and are sufficient to ensure, only that evidential sentences be 'pragmatically decidable' and 'non-problematic', in senses that will be explained shortly. We shall have no formal need of the notion of *certainty* in any of its classificatory, comparative, or quantitative forms, though we shall make use of it informally in discussing the intuitive adequacy of our axioms.

The property of being an evidential sentence will be relativised here to individual persons. If, at some time, each member of a certain community is led to the same class of evidential sentences, the sentences of this class may be called evidential for the community at that time. In making an extension to communities we might profitably use the notion of 'inter-subjective agreement', but we shall make do here without it.

We have been led in the last two sections to a system of axioms for acceptance which are to hold for all members of S_L (the class of all sentences of the object-language L). But for some sub-classes of S_L —the class of evidential sentences, for example—stronger conditions may be satisfied. This is the idea behind the axioms for the new primitive predicate ' E '. ' $E(e)$ ' is intended to be read as ' e is an evidential sentence for X at time t '. (' E ', like ' A ', is understood to make implicit reference to a person and a time.) ' E ' is introduced in the following axioms:

$$(E1) \quad E(e) \rightarrow (\sim A(\sim e) \rightarrow A(e)),$$

$$(E2) \quad \vdash (e \leftrightarrow f) \rightarrow (E(e) \leftrightarrow E(f)),$$

$$(E3) \quad E(e) \& E(f) \rightarrow E(e \& f),$$

$$(E4) \quad E(e) \rightarrow E(\sim e).$$

Before proceeding to a discussion of the intended interpretation of ' E ', we note first that the consistency of ($\alpha A1-3$) and ($E1-4$), relative to the theories of syntax and semantics presupposed, can be established by taking ' A ' to mean 'is true' (in the semantic sense) and ' E ' to mean 'is a sentence', other symbols being given their usual interpretations.

Secondly, we note the following consequences of ($\alpha A1-3$) and ($E1-4$):

$$E(e) \rightarrow (A(e) \leftrightarrow \sim A(\sim e)), \quad (27)$$

$$E(e) \& E(f) \rightarrow (A(e \vee f) \leftrightarrow A(e) \vee A(f)), \quad (28)$$

$$E(e) \& E(f) \rightarrow (A(e \& f) \leftrightarrow A(e) \& A(f)), \quad (29)$$

$$E(e) \& E(f) \rightarrow (A(e \rightarrow f) \leftrightarrow (A(e) \rightarrow A(f))). \quad (30)$$

These play an important part in the following discussion of the interpretation of 'E'.

I shall suppose that the class of evidential sentences is somehow given. I shall try to justify the interpretation of 'E' as the property of being an evidential sentence by listing some characteristics of sentences with the property *E* in any given model—for the moment to be called '*E*-sentences'—which also seem plausible characteristics of evidential sentences.

First, no *E*-sentence may be left in suspense. Suspension of judgment was defined above as holding for sentences satisfying ' $\sim A(e) \& \sim A(\sim e)$ ', and this possibility is ruled out by (E1). *E*-sentences may therefore be termed 'decidable in the pragmatic sense'.

Secondly, and similarly, (E1-4), together with (α A1-3), prevent *X* from remaining in doubt between *E*-sentences in the following sense. *X* will be said to *remain in doubt between* the sentences $e_1, e_2, \dots, e_n$ if and only if he accepts the disjunction $e_1 \vee e_2 \vee \dots \vee e_n$ but accepts none of the disjuncts individually, *or* rejects the conjunction $e_1 \& e_2 \& \dots \& e_n$ while rejecting none of the conjuncts individually. Now from (E2-4) we have

$$E(e) \& E(f) \rightarrow E(e \vee f).$$

(28) may therefore be extended to any finite disjunction as

$$E(e_1) \& E(e_2) \& \dots \& E(e_n) \rightarrow (A(e_1 \vee e_2 \vee \dots \vee e_n) \leftrightarrow A(e_1) \vee A(e_2) \vee \dots \vee A(e_n)). \quad (31)$$

Thus if *X* accepts any finite disjunction of *E*-sentences, he must also accept at least one of the disjuncts. It follows similarly that if he rejects a finite conjunction of *E*-sentences, he must also reject at least one of the conjuncts. Hence he cannot remain in doubt between the members of a set of *E*-sentences. This may be put in another way using David Harrah's logic of questions (see Harrah, 1963, especially pp. 31-2). It follows from (31), (α A1), (α A2) that if *X* accepts a *disjunctive question* in some sequence of *E*-sentences he must also accept one and only one *direct answer* to the question. *E*-sentences may therefore be called 'non-problematic'.

Thirdly, the class of *accepted E*-sentences (denoted by 'AE') is closed under the operation of finite conjunction; i.e. any finite conjunction of accepted *E*-sentences is likewise an accepted *E*-sentence. For (E3) may be extended to

$$E(e_1) \& E(e_2) \& \dots \& E(e_n) \rightarrow E(e_1 \& e_2 \& \dots \& e_n), \quad (32)$$

and, by repeated applications of (E3), (29) can be extended to

$$E(e_1) \& E(e_2) \& \dots \& E(e_n) \rightarrow (A(e_1 \& e_2 \& \dots \& e_n) \leftrightarrow A(e_1) \& A(e_2) \& \dots \& A(e_n)). \quad (33)$$

Then, by (32) and (33), we have the required result

$$AE(e_1) \& AE(e_2) \& \dots \& AE(e_n) \rightarrow AE(e_1 \& e_2 \& \dots \& e_n), \quad (34)$$

where 'AE' is a new abbreviation defined by

$$AE(e) \leftrightarrow A(e) \& E(e)$$

Fourthly, it follows from (34) and ($\alpha A2$) that every logical consequence of a set of accepted *E*-sentences is likewise accepted.

Fifthly, the class **AE** of accepted *E*-sentences is logically consistent. Suppose **AE** is non-empty (otherwise it is trivially consistent). If then it is inconsistent, it must contain a finite subclass of sentences the conjunction of whose members, *c* say, is logically false. But by (34), *c*, being a conjunction of accepted *E*-sentences, must be accepted. Since **AE** is non-empty by hypothesis, at least one sentence of *L* is accepted; thus by (14) every logically true sentence, in particular $\sim c$, is accepted. Since both *c* and $\sim c$ are accepted we have a contradiction with ($\alpha A1$). **AE** must therefore be consistent.

Sixthly, if **AE** is non-empty it contains all logical truths of *L*. We have seen that if **AE** is non-empty every logical truth of *L* is accepted. It remains to show that each such sentence is an *E*-sentence. If **AE** is non-empty there must be at least one *E*-sentence, *e* say. In that case, by (E2-4), $e \vee \sim e$ is also an *E*-sentence; hence by (E2) every logical truth of *L* is an *E*-sentence, and the proof is complete. It follows further, with (E4), that every logically determinate sentence of *L* is an *E*-sentence.

It seems reasonable to expect evidential sentences to possess these six characteristics of *E*-sentences. (The results we have obtained for *accepted E*-sentences are among those required by Isaac Levi to hold for sentences 'accepted as evidence'; (see Levi, 1967, pp. 28-29.) But our axioms will be inappropriate if they make further demands of *E*-sentences which cannot be made of evidential sentences. In this respect (E4) deserves attention. For if 'This pen is red' is an *E*-sentence, so also, by (E4), is 'This pen is not red'. But can the latter be called evidential? If evidential sentences are identified with reports of 'direct observation', the answer must be 'No'. For one cannot *directly* observe something *not* to be red. One must first observe it to be green, for example, and then deduce its not being red with the help of a linguistic convention or some known 'principle of necessity'. But in this trivial deduction there is no diminution in certainty between premiss and conclusion. If we are certain that a thing is green, we are equally certain that it is not red; if we know it to be green, we know it not to be red. On our understanding of evidential sentences this is sufficient reason for including the second sentence mentioned above into the class of evidential sentences. But the presence of (E4) among our axioms

indicates one way at least in which we are committed to a less restrictive view of the nature of evidence than has been proposed in some other accounts. We shall see how our account is less restrictive in another more important sense shortly.

There is a further point in favour of the legitimacy of our intended interpretation of 'E'. Evidential sentences might well be understood as those attributed probabilities of 0 or 1 by *X* at time *t*. Such sentences will be called *probabilistically determinate* (for *X* at time *t*). Then if *X* satisfies (PRA), it can easily be shown consistent to interpret 'E' as the property of being probabilistically determinate, and therefore as that of being an evidential sentence on this understanding.

On the other hand, even if (PRA) is satisfied, (E1-4) do not, by themselves, require *E*-sentences to have probabilities of 0 or 1. Nevertheless they place severe restrictions on these probabilities. Roughly speaking, the larger the number of logically distinct *E*-sentences that are not probabilistically determinate, the nearer these sentences must approach determinacy, i.e. probabilities of 0 or 1.¹

This point has some bearing on the sufficiency of (E1-4) as a characterisation of evidential sentences and consequently on the way in which the class of such sentences is to be defined. Two suggestions will be made in this connection.

Given the interpretation of 'A' as meaning '(X) accepts (at time *t*)' in some appropriate sense of 'accepts', any class of sentences that may be taken, consistently with our axioms, as the extension of 'E' will be called an *E-type class*.

Bearing in mind that the defining properties of *E*-type classes are not sufficient to ensure that the members of such classes are even probabilistically determinate, from an epistemological point of view there seem to be two possibilities open. Either (i) we say that the class of evidential sentences is to be an *E*-type class but that further conditions must be included in any adequate specification of it, ensuring, in particular, that the truth-values of its members are known with certainty; or (ii) we say that any *E*-type class may be taken as the class of evidential sentences, and that the characteristics of such classes impose *sufficient* restrictions on the *degrees* of certainty with which the truth-values of evidential

¹ John Maynard Keynes has required the sentences that are to stand as evidence for rational beliefs to be *known* (Keynes, 1921, p. 16). If they may legitimately be attributed probabilities, it would appear that, for Keynes, they must have 'the highest degree of rational belief, which is termed *certain* rational belief' (p. 10), i.e. they must have probabilities of 1. The certainty with which accepted evidential sentences should be known has been discussed more recently, for example, by Barker (1957, pp. 31-43), Goodman (1952, especially pp. 162-3), and Chisholm (1957, p. 25).

sentences, now defined as the members of the particular *E*-type class chosen, are known.

In the case of (i), one further condition might be the requirement that each member of the class of accepted evidential sentences, other than those that are logically determinate, must be an element of the counterdomain of an *empirical meaning-rule* in the following sense of Ajdukiewicz (1949, p. 183):

This kind of meaning-rule demands the readiness to accept certain sentences in the presence of certain data of experience. I call such rules *empirical meaning-rules*, as by them are specified sentences which can be established with certainty, in a purely empirical way.

(i) would then be consistent with the approach adopted here of analysing the concept of evidence as far as possible in terms of acceptance alone.¹

(ii) is also consistent with this approach but goes much further towards completing the programme suggested by it. (ii), however, allows a large element of choice in the specification of the class of evidential sentences, since there will be very many *E*-type subclasses of S_L from which to choose. Suppose, for example, that *h* is any sentence about which judgment is not suspended. Then the class of sentences comprising just *h*, $\sim h$, $h \& \sim h$, $\sim(h \& \sim h)$, and every sentence of *L* logically equivalent to one of these, is an *E*-type class. But in this way, many *E*-type classes will be wholly inadequate for the purposes of science. Some criterion of maximality, relative to the problem currently under investigation by the scientist, might be used to remove the arbitrariness in (ii). But such a criterion would arise from methodological, and not epistemological, considerations. According to (ii), any *E*-type class is satisfactory from an epistemological point of view, in possessing the requisite structure with respect to acceptance.

In conclusion, it is of some interest to note how our analysis of evidential sentences allows these to be, to a certain extent, theoretical. According to (ii), and also to (i) if appropriate further conditions are employed, it is not necessary for evidential sentences to be purely 'observational' or even 'known with certainty'. It is sufficient that none of them individually, nor any truth-functional combination of them jointly, is left in suspense. This means that it will be possible to take sentences as evidential whose acceptance or rejection, and even whose comprehensibility, depend crucially on scientific theories, if the relevant theories are sufficiently well accepted, and the meanings of their terms sufficiently familiar, for these sentences to be accepted or rejected in accordance with the patterns of acceptance holding over *E*-type classes. In this way sentences which are

¹ For an analysis of *meaning* primarily in terms of *acceptance*, see Ajdukiewicz (1934).

at one time hypothetical may later become evidential as related theories become more and more firmly established.

Our analysis of the structure of acceptance, though shedding light on only a fragment of the multiplicity of senses in which this concept is used among philosophers and scientists, has therefore enabled us to avoid the notion of 'observationality', and, from a formal point of view, that of 'certainty', in the discussion of evidence. In place of an analysis based on an understanding of these elusive concepts, we have been able to provide a less restrictive pragmatic account presupposing only a classificatory concept of acceptance and its logical structure.

The University of Sussex

REFERENCES

- AJDUKIEWICZ, K. (1934). Sprache und Sinn, *Erkenntnis* 4, 100-38.
 AJDUKIEWICZ, K. (1949). The Scientific World Perspective. In *Readings in Philosophical Analysis*. Ed. H. Feigl and W. Sellars, New York, pp. 182-8.
 BARKER, S. F. (1957). *Induction and Hypothesis*, Ithaca.
 CARNAP, R. (1962). The Aim of Inductive Logic. In *Proc. 1960 Intern. Congress for Logic, Methodology and Philosophy of Science*. Ed. E. Nagel, P. Suppes and A. Tarski, Stanford, pp. 303-18.
 CARNAP, R. (1964). Replies and Systematic Expositions. In *The Philosophy of Rudolf Carnap*. Ed. P. A. Schilpp, La Salle, pp. 859-1013.
 CHISHOLM, R. M. (1957). *Perceiving: A Philosophical Study*. Ithaca.
 DE FINETTI, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré* 7, 1-68.
 GOODMAN, N. (1952). Sense and Certainty. *Phil. Rev.* 61, 160-7.
 HARRAH, D. (1963). *Communication: A Logical Model*. Cambridge, Mass.
 HEMPEL, C. G. (1962). Deductive-Nomological vs. Statistical Explanation. In *Minnesota Studies in the Philosophy of Science*, Vol. 3. Ed. H. Feigl and G. Maxwell. Minneapolis, pp. 98-169.
 HINTIKKA, J. & HILPINEN, R. (1966). Knowledge, Acceptance, and Inductive Logic. In *Aspects of Inductive Logic*. Ed. J. Hintikka and P. Suppes. Amsterdam, pp. 1-20.
 JEFFREY, R. C. (1956). Valuation and Acceptance of Scientific Hypotheses. *Phil. of Sci.* 23, 237-48.
 KEYNES, J. M. (1921). *A Treatise on Probability*. London.
 KYBURG, H. E. (1965). Probability, Rationality, and a Rule of Detachment. In *Proc. 1964 Intern. Congress for Logic, Methodology, and Philosophy of Science*. Ed. Y. Bar-Hillel, Amsterdam, pp. 301-10.
 LEVI, I. (1960). Must the Scientist make Value Judgements? *J. Phil.* 57, 345-57.
 LEVI, I. (1967). *Gambling with Truth*. New York and London.
 MARTIN, R. M. (1959). *Toward a Systematic Pragmatics*. Amsterdam.
 PRIOR, A. N. (1962). *Formal Logic*, 2nd edn., Oxford.
 RUDNER, R. (1953). The Scientist *Qua* Scientist makes Value Judgements. *Phil. Sci.* 20, 1-6.
 SLEIGH, R. C. (1964). A Note on some Epistemic Principles of Chisholm and Martin. *J. Phil.* 61, 216-8.
 VON WRIGHT, G. H. (1951). *An Essay in Modal Logic*. Amsterdam.
 VON WRIGHT, G. H. (1963). *Norm and Action*. London.

Discussions

NEW MYSTERIES FOR OLD: THE TRANSFIGURATION OF MILLER'S PARADOX

Let A be the statement that a certain neutrally identified object has property a ;¹ let ' $P(A, X)$ ' denote the probability (credibility) that A is true given information X ; and let ' E_r^a ' abbreviate the statement ' $fr(a)=r$ ', where $fr(a)$ is the statistical frequency (or statistical probability, or observed frequency) of property a . It is then strongly intuitive to think that for any r ,

$$(1) \quad P(A, E_r^a) = r,$$

i.e. to suppose that the probability of an object's having property a , given only the information that property a has statistical frequency r , is likewise equal to r . However, Miller (1966) and Popper (1966) have recently argued that unqualified acceptance of (1) leads to paradox, and subsequent protests by Mackie (1966) and Bub and Radner (1968) have not diminished the cogency of Miller's argument (cf. Miller 1966a, 1968). Indeed, they could not; for the argument is entirely sound. The *significance* of Miller's paradox, however, is rather different from what it seems: It does not in any way speak against (1) as a principle of inductive logic, but merely emphasises an important boundary condition which must be imposed on *any* inference schema involving hypotheticals. Even so, while Miller's own version of his paradox is indifferent to the specific content of (1), there are indeed reasons why our intuitive acceptance of the latter should not be unhesitant. For substantive paradox does in fact lurk within (1), and the spirit of this paradox remains an ominous presence even after it has been technically exorcised by suitable restrictions on (1)'s generality.

1

To develop Miller's paradox from generalisation (1), consider the hypothesis that $fr(a)=fr(\sim a)$, i.e. that $E_{fr(\sim a)}^a$ where $\sim a$ is the property of not possessing property a . It is easily proved that $fr(a)=fr(\sim a)$ if and only if $fr(a)=.5$; hence the statements $E_{fr(\sim a)}^a$ and $E_{.5}^a$ are logically equivalent. Moreover, for any two

¹ Stipulation that statement A 'neutrally identifies' the object of which it predicates a is intended to exclude cases of the sort illustrated by a 's being the property 'is taller than John' while A is 'John is taller than John'. That is, we must forbid any special analytic connection between the predicate ' a ' and the subject term to which a is ascribed in A .

logically equivalent statements B and B' , $P(A, B)=P(A, B')$. Consequently, were (1) to hold with unlimited generality, it would follow that

$$(2) \quad fr(\sim a)=P(A, E_{fr(\sim a)}^a)=P(A, E_{.5}^a)=.5,$$

which paradoxically constrains $fr(\sim a)$ and hence $fr(a)$ to have a numerical value of .5 when in fact $fr(a)$ is a free parameter. Unless we abandon the interchangeability of analytically equivalent statements in probability contexts, this argument cannot be faulted. Principle (1) is simply not acceptable *with unrestricted substitutability for its free variables*. It still remains to see, however, whether the needed exclusions are specific to the content of (1) or are merely technical qualifications binding on all principles of propositional probability, akin to forbidding division of zero by zero in otherwise universal theorems of algebra.

For example, one technically essential but conceptually trivial boundary restriction on (1) is evident in the following argument: Let u be the property of self-identity—i.e. for any entity x , $u(x)$ iff $x=x$ —and let r and s be any two numbers different both from each other and from unity. Since it is logically true that $fr(u)=1$, E_r^u and E_s^u are both logically false and hence logically equivalent to each other. (In fact, if r and s are any entities other than numbers in the unit interval, this will be true regardless of what property u may be.) Consequently, from (1)

$$r=P(A, E_r^u)=P(A, E_s^u)=s,$$

contrary to stipulation that $r \neq s$. The source of this absurdity is allowing (1) to subsume instances in which E_r^a is logically false. That trouble should arise from this is not, however, at all peculiar to (1). For under any standard axiomatisation of propositional probability, it is elementary to show that for any two propositions A and B such that B entails A , $P(A, B)=1$ and $P(\sim A, B)=0$ so long as the quantity $P(A, B)$ is well-defined. Hence if $P(A, F)$ were to be well-defined for some logically false proposition F and another proposition A , it would inconsistently follow, since F then entails both A and $\sim A$, that $1=P(\sim A, F)=0$. Consequently, any consistent theory of propositional probabilities must treat as ill-defined all conditional probabilities in which the conditional's hypothesis is logically false. No matter how universally (1) may otherwise be conjectured to hold, therefore, we must in any event stipulate that it presupposes a logically consistent protasis E_r^a . Henceforth, I shall refer to this as the 'consistency requirement'.

At first impression, Miller's paradox seems to be a violation of the consistency requirement; for the argument involves hypothesising that $fr(a)=fr(\sim a)$ when we have also stipulated that $fr(a)$ has some value other than .5. But while the assertion ' $fr(a)=fr(\sim a)$ & $fr(a) \neq .5$ ' is indeed inconsistent, this conjunction does not figure in the formal deduction. Unwanted conclusion (2) follows merely by applying (1) to $E_{fr(\sim a)}^a$ and $E_{.5}^a$ separately, and each of these protases is in

itself logically consistent (assuming a to be an empirical property) even if in fact $fr(a)$ has some value other than .5. To avert (2) we thus require a stronger restriction on (1) than just the consistency requirement.

To appreciate what this further restriction must be, consider first of all the fallacy in the following argument:

If Oswald's hand had trembled that moment in Dallas,
J. F. Kennedy would have been the U. S. President in 1964.
The U.S. President in 1964 was L. B. Johnson

[Therefore], if Oswald's hand had trembled that moment in Dallas,
J. F. Kennedy would have been L. B. Johnson.

Syntactically, this inference may be parsed according to the schema

$$\frac{F(x)}{x=y} \\ \frac{}{F(y)}$$

which is formally valid. However, *formally* valid inference schemata, when instantiated, yield *logically* valid inferences only when each free syntactic variable is replaced by a term which maintains the same referent throughout the argument.¹ This rule is deliberately violated in the present case by construing the descriptor 'the U.S. President' to have a context-dependent referent which varies from one premise to another. While considerably more could be said about this example, it suffices to demonstrate that formally valid arguments in which hypothetical premises are involved may well prove logically treacherous if these include descriptors.

Once one stops to think about it, it is evident that Miller's paradox results from precisely this sort of referential ambivalence. For inasmuch as $\vdash E_{fr(\sim a)}^a$

$\equiv E_{.5}^a$, the statement

$$(3) \quad P(A, E_{fr(\sim a)}^a) = fr(\sim a)$$

is logically equivalent to

$$(4) \quad P(A, E_{.5}^a) = fr(\sim a),$$

which is hence a consequence of (1) iff (3) is. Yet (4) is not at all the sort of conclusion which principle (1) is intended to authorise. What has gone wrong here is that the rightmost occurrence of ' $fr(\sim a)$ ' in (3) and (4) is construed to designate whatever number is *in fact* the value of $fr(\sim a)$ (or, more technically, whatever value is stipulated for this parameter elsewhere in the argument), with the result that (3) asserts that the probability of A , given that $fr(a) = fr(\sim a)$, equals the actual frequency of $\sim a$. But what (1) says is that the probability of A , given that $fr(a)$ equals some particular number, has the same numerical value as the one stipulated for $fr(a)$ by the hypothesis. Thus while (3) is formally an instance of (1), its descriptor ambivalence disqualifies it as a *logical* consequence of the latter.

¹ For discussion of the distinction between formal validity and logical validity, see Rozeboom, 1962, p. 17.

It may be noted, incidently, that the equivalence of ' $fr(\sim a) = .5$ ' with

$$(5) \quad fr(a) = fr(\sim a)$$

requires that (5) be interpreted as

$$(5') \quad \text{Hypothesis: } fr(a) \text{ and } fr(\sim a) \text{ have the same value.}$$

But while (5') is the most natural reading of (5), a perfectly legitimate alternative interpretation is

(5'') *Hypothesis: $fr(a)$ equals that number which is in fact the value of $fr(\sim a)$.* That (5') and (5'') are very different hypotheses may be appreciated by reflecting that (5') entails that $fr(\sim a) = .5$ whereas (5'') does not. Rather, combined with the auxiliary hypothesis, say, that $fr(\sim a)$ in fact equals .2, (5'') becomes the hypothesis that $fr(a) = .2$.

Without attempting to do full justice to the complexities of the situation here revealed, we may say that each occurrence of a descriptor generally carries a rider—usually implicit in the context of usage—resolving uncertainty about the descriptor's intended referent. In most indicative contexts, a descriptor d designates whatever entity in fact satisfies d . In subjunctive or counterfactual contexts, however, d 's referent is often construed to be whatever it *would* be were some state of affairs s to be the case, or what it would have to be in order for a certain sentence $S(d)$ incorporating d to be true. Consequently, two different occurrences of the same descriptor in a hypothetical argument cannot be treated *logically* as the term, despite their formal (i.e. syntactic) identity, unless they carry the same reference rider. In particular, formally valid deductions from any conditional-probability generalisation may well eventuate in logical fallacy unless the restriction is imposed that no instantiation of a free (or universally quantified) variable is allowed by an expression whose referent is context dependent. We may call this the 'univalence requirement'. In no way does the univalence requirement diminish a generalisation's universality; it merely forbids certain syntactic manoeuvres which are valid formally but not logically.

2

So long as descriptor ' $fr(\sim a)$ ' is allowed to have different reference riders in different contexts of occurrence, (3) cannot be deduced from (1) without violating the univalence requirement; hence this restriction thwarts Miller's paradox. (To be sure, (1) still entails (3) when the latter is read as 'The probability of A , given that $fr(a) = fr(\sim a)$, equals the value that $fr(\sim a)$ would have if $fr(a) = fr(\sim a)$ '; but then conclusion (2) correspondingly reads ' $.5$ is the value that $fr(\sim a)$ would have if $fr(a) = fr(\sim a)$ ', which is hardly paradoxical.) However, it still remains to see whether the restrictions we have found necessary to impose on (1) so far also cut the ground from under Popper's two variants of the paradox. One of Popper's versions (1966, p. 65) is essentially the same as Miller's and is similarly defeated by the univalence requirement. Popper's main version (1966, p. 64), however, cannot be disposed of quite this briskly. Essentially in Popper's own words, the argument goes as follows:

For any x whatever,

(6) If $\vdash s = 2r - x$, then the following equivalence holds analytically:

$$x = r \text{ if and only if } x = s.$$

By substituting ' $fr(a)$ ' for ' x ' we obtain from (6)

(7) If $\vdash s=2r-fr(a)$, then the following equivalence holds analytically:
 $fr(a)=r$ if and only if $fr(a)=s$.

By using the metalinguistic names ' E_r^a ' and ' E_s^a ', this can be written

(8) If $\vdash s=2r-fr(a)$, then $\vdash E_r^a \equiv E_s^a$.

But (8) allows us to substitute ' E_r^a ' and ' E_s^a ' in every probability formula for each other. Thus from (1) we obtain

(9) If $\vdash s=2r-fr(a)$, then $P(A, E_r^a) = P(A, E_s^a) = r = s$

and therefore

(10) If $\vdash s=2r-fr(a)$, then $r=s$.

But this result, it can be seen at once, is inconsistent; for we may for any given number, $fr(a)$, choose an r and an s such that r is not equal to s even though the conditions of (10) are satisfied.

At no point does the wording of this argument manifestly violate the consistency or univalence restriction on (1). It does, however, subtly equivocate between two different readings of the symbol ' $fr(a)$ ', and when this ambiguity is resolved the argument fails in one of three ways. Observe to begin with that step (7), which on first impression seems to assert that certain statements are logically equivalent if a certain other statement is logically true, is in fact only the *schema* of such an assertion and becomes a genuine proposition only when the metalinguistic variables ' s ', ' r ', and ' $fr(a)$ ' are replaced by actual expressions in the object-language.¹ When this schema is instantiated, ' $fr(a)$ ' can be replaced either by (i) a *numeral*, or by (ii) a descriptive phrase synonymous with 'the frequency of a '. If (i), the transition from step (7) to step (8) is invalid. But if (ii), then either the antecedent of the step (7) conditional is false or the consistency restriction on (1) blocks passage from (8) to (9), so that again no paradox is obtained. A specific example of each alternative will show how this is so:

For case (i), let ' s ', ' r ' and ' $fr(a)$ ' be instantiated by ' $\cdot 8$ ', ' $\cdot 4$ ', and ' $\cdot 6$ ', respectively. This converts schema (7) into the determinate theorem

(7') If $\vdash \cdot 8 = 2 \times \cdot 6 - \cdot 4$, then the following equivalence holds analytically:
 $\cdot 4 = \cdot 6$ if and only if $\cdot 4 = \cdot 8$

Observe however, that the apodosis of this conditional, namely, that $\vdash \cdot 4 = \cdot 6 \equiv \cdot 4 = \cdot 8$, is *not* equivalent to the assertion that $\vdash E_{\cdot 6}^a \equiv E_{\cdot 8}^a$; for the latter abbreviates $\vdash fr(a) = \cdot 6 \equiv fr(a) = \cdot 8$, in which ' $fr(a)$ ' is not a metalinguistic variable but a synonym for the object-language descriptor 'the frequency of a '. The antecedent and hence also the consequent of conditional (7') are true but do *not* entail the consequent of the corresponding instantiation of (8), namely, that 'the frequency of a is $\cdot 6$ ' is analytically equivalent to 'the frequency of a is $\cdot 8$ '.

Alternatively—case (ii)—let ' s ' and ' r ' be respectively instantiated by ' $\cdot 8$ ' and ' $\cdot 6$ ' as before, but this time let ' $fr(a)$ ' be read as a synonym for object-

¹ Or, more precisely, by metalinguistic designators of such expressions. A certain amount of harmless waffling between use and mention is inescapable here unless we resort to tediously complicated formulations.

language descriptor 'the frequency of a '. Schema (7) then becomes the determinate theorem

(7'') If $\vdash .8 = 2x \cdot 6 - fr(a)$, then the following equivalence holds analytically:
 $fr(a) = .6$ if and only if $fr(a) = .8$.

This is indeed equivalent to

(8'') If $\vdash .8 = 2x \cdot 6 - fr(a)$, then $\vdash E_{.6}^a \equiv E_{.8}^a$.

from which together with (1) we may conclude, so long as neither $E_{.6}^a$ nor $E_{.8}^a$ are logically false,

(9'') If $\vdash .8 = 2x \cdot 6 - fr(a)$, then $P(A, E_{.6}^a) = P(A, E_{.8}^a) = .6 = .8$.

The paradoxical conclusion that $.6 = .8$, however, is conditional upon its being the case that $\vdash .8 = 2x \cdot 6 - fr(a)$, i.e. upon its being logically true that $fr(a) = .4$.

But if the latter is the case, then both $E_{.6}^a$ and $E_{.8}^a$ are logically false, whence (9'') no longer follows from (8'') and (1) under the consistency restriction on (1).

There are still other ways in which schema (7) might be instantiated, notably by letting ' s ' and ' r ' be descriptors, but in all cases derivation of the paradox is thwarted either by counterfactuality of the resultant conditional or, if the instantiated $\vdash s = 2r - fr(a)$ is in fact logically true, by the inconsistency of E_r^a or

E_s^a . We may conclude, then, that unlike Miller's simple but effective own version of his paradox, Popper's elaboration is merely a demonstration of subtly invalid inference.

3

Having defended the substantive generality of (1) against previous objections, I shall now act turncoat—reluctantly, since (1) still has considerable intuitive appeal to me—and raise some qualms about its specific content.

Principle (1) is one possible answer to the still unsolved question of how propositional probabilities relate to statistical frequencies. Another is proposed by what, for want of a better name, may be called

The Thesis of Statistical Reduction: If sentence ' $a(x)$ ' asserts that entity x has property a , where term ' x ' has no special analytic connection with predicate ' a ' (see fn. 1), then the unconditional probability of $a(x)$ is $P[a(x)] = fr(a)$.

(There are actually as many versions of this *Thesis* as there are interpretations for the functor ' fr ', notably, frequency, statistical probability, or observed frequency. We assume that ' fr ' is to be given the same reading here as in (1), but otherwise its specific interpretation may be left open.)

Although the Thesis of Statistical Reduction seems intuitively more dubious than does principle (1), a slight modification of Miller's paradox shows that (1) actually entails it. For since ' $a(x)$ ' may be taken as the ' A ' in (1), substituting ' $fr(a)$ ' for ' r ' in (1) yields

$$(11) \quad P[a(x), E_{fr(a)}^a] = fr(a).$$

But ' $E_{fr(a)}^a$ ' simply abbreviates ' $fr(a) = fr(a)$ ', which is tautologous. Inasmuch as the conditional probability of any proposition given a tautology equals its unconditional probability, (11) is thus equivalent to

$$(11') \quad P[a(x)] = fr(a).$$

Hence (1) entails (11').

Unhappily, however, the Thesis of Statistical Reduction suffers from a fatal flaw when asserted in full generality. For let sentence ' $b(x, y)$ ' assert that entity x stands in relation b to entity y —e.g. 'John loves Mary', or 'Peter beats Michael the first time they play chess together'—and let ' b_{x*} ' and ' b_{*y} ' abbreviate the monadic predicates ' $b(x, —)$ ' and ' $b(—, y)$ ' respectively. Since by definition,

$$(12) \quad \vdash b_{x*}(y) \equiv b(x, y) \equiv b_{*y}(x),$$

while from (11'),

$$(13a) \quad P[b_{x*}(y)] = fr(b_{x*})$$

$$(13b) \quad P[b(x, y)] = fr(b)$$

$$(13c) \quad P[b_{*y}(x)] = fr(b_{*y}),$$

it follows by the principle that logically equivalent propositions have equal unconditional probabilities that

$$(14) \quad fr(b_{x*}) = fr(b) = fr(b_{*y}).$$

But it is just not true that (14) is generally the case. The frequency of persons loved by John, for example, has no analytic nor even nomic bearing on the frequency of persons who love Mary; while if Peter and Michael are both good chess players, the frequency of 'Peter beats — the first time they play chess together' is high, the frequency of '— beats Michael [etc.]' is low, and the frequency of '— beats — [etc.]' is .5 less the frequency of draws.

If principle (1) is to be retained at all, therefore, it must be curtailed even more severely than is accomplished by the consistency and univalence requirements.¹ Precisely what restriction is most germane is not altogether evident.

¹ Or does perhaps the univalence requirement forbid inference of (11) from (1)? Since ' $E_{fr(a)}^a$ ' is tautologous no matter what reference rider is attached to ' $fr(a)$ ' therein (so long as it is the same for both occurrences), it would seem that we can construe this rider to be the same as the one holding for the rightmost occurrence of ' $fr(a)$ ' in (11), thereby satisfying the univalence requirement. On the other hand, it might conversely be argued that occurrence of a descriptor in the protasis of *any* conditional gives it a somewhat different meaning from its occurrence outside of that hypothesis, so that (11) is not a permissible instance of (1). The fact that this point is not entirely clear shows that my previous discussion of the univalence requirement is incomplete, and that the functioning of descriptors in hypothetical inference still lacks a fully satisfactory analysis. (*Retraction added in proof*: It now seems clear to me that the univocality requirement does indeed forbid inference of (11) from (1), or more precisely, that it authorizes the move to (11) only when $E_{fr(a)}^a$ is construed in a sense which is not tautological. Oh, well; the argument from (1) to (11') is intriguing even if invalid, and its failure does no harm to the remaining argument.)

Our primary intuitive understanding of (1), however, is that its open variable ' r ' is to be replaced by a specific numeral or at least by an expression whose referent can be identified by us. (Instances of (1) just don't *feel* quite right when the r -term designates an unknown, e.g. 'the proportion of 23rd Century astronauts with blonde hair'.) As a working hypothesis, therefore, let us provisionally impose on (1) the 'determinateness' requirement that its universality is to cover only numerical instances, i.e. that a conditional probability assertion of form (1) is authorised by this principle only when ' r ' is, or is analytically equivalent to, a specific numeral. Under this determinateness requirement, (11) is no longer an authorised instance of (1), so paradox (14) is thereby averted. (Note, however, that this does not salvage the Thesis of Statistical Reduction; it merely prevents (1) from entailing it. What might be done to resuscitate principle (11') will not be discussed here.)

Even so, the considerations underlying paradox (14) continue to haunt principle (1) no matter how severely the latter's generality is restricted. Let b_{x^*} and b_{x^*y} be defined from binary relation b as before, so that (12) continues to hold. Then what is the conditional probability of $b(x,y)$ given both that $fr(b_{x^*}) = r$ and that $fr(b_{x^*y}) = s$, where ' r ' and ' s ' are specific numerals such that $r \neq s$? For example, what is the probability that Peter beats Michael the first time they play chess together, given that Peter wins 60 per cent of his first chess games with other players while Michael is beaten on 10 per cent of his first chess games with others? By principle (1), the probability is .60 that Peter beats Michael given only that Peter wins 60 per cent of his games, whereas the probability is .10 that Peter beats Michael given only that Michael is beaten on 10 per cent of his games. There is no inconsistency in these probabilities, for each is conditional on a different statistical hypothesis. But these two statistics are entirely compatible, and if they are given jointly, *then* what is the conditional probability that Peter beats Michael? The answer—or rather, non-answer—is that intuition simply breaks down in this case. To be sure, we can think of possibilities, such as

$$(15) \quad P[b(x,y), E_r^{b_{x^*}} \& E_s^{b_{x^*y}}] = .5(r+s),$$

but (15) is in general false (cf. the case where r or s is zero or unity) and there are many alternatives to (15) which might be entertained, each as intuitively arbitrary as the others.

Insomuch as there exists no general principle which, for any three propositions A, B , and C , derives $P(A, B \& C)$ from $P(A, B)$ and $P(A, C)$, there is no reason why (1) should be held responsible for providing the numerical value of $P[b(x,y), E_r^{b_{x^*}} \& E_s^{b_{x^*y}}]$. The plausibility of (1), however, is now seen to depend on the strength of one's conviction that there exists a numerical function f such that for any two numerals r and s ,

$$(16) \quad P[b(x,y), E_r^{b_{x^*}} \& E_s^{b_{x^*y}}] = f(r, s).$$

For if there is no such f , i.e. if the specific numerical frequencies of $b(x, —)$ and $b(—, y)$ do not suffice to determine the conditional probability of $b(x,y)$ given these frequencies, then whatever additional factors influence this probability

may well be expected to make a difference for $P[a(x), E_r^a]$ as well, thereby preventing the latter from being determined solely by frequency datum E_r^a as alleged by (1). Personally, I consider it fairly plausible that, under suitable boundary restrictions, generalisation (16) is in fact true for some f . But until this function has actually been identified and submitted to probate, (16) and hence (1) must remain an article of more or less tenuous faith.

WILLIAM W. ROZEBOOM
University of Alberta

REFERENCES

- BUB, J. and RADNER, M. (1968). Miller's paradox of information. *Brit. J. Phil. Sci.* 19, 63-7.
 MACKIE, J. L. (1966). Miller's so-called paradox of information. *Brit. J. Phil. Sci.* 17, 144-7.
 MILLER, D. (1966). A paradox of information. *Brit. J. Phil. Sci.* 17, 59-61.
 MILLER, D. (1966a). On a so-called paradox: A reply to Professor J. L. Mackie. *Brit. J. Phil. Sci.* 17, 147-9.
 MILLER, D. (1968). The straight and narrow rule of induction. *Brit. J. Phil. Sci.* 19, 145-51.
 POPPER, K. R. (1966). A comment on Miller's new paradox of information. *Brit. J. Phil. Sci.* 1966, 17, 61-9.
 ROZEBOOM, W. W. (1962). Intentionality and existence. *Mind*, 71, 15-32.

NEUTRALITY, CONTINGENCY AND UNDECIDABILITY

Professor Körner makes no reference to his 'modified two-valued logic' or to his thesis of 'free choice' between truth and falsehood for each neutral atomic proposition in indicating his answers (this *Journal*, February 1968, briefly, *R*) to my criticisms (this *Journal* November 1967) of his views regarding 'the logic underlying empirical discourse'. And he admits that 'the three-valued logic . . . contains no forms all substitution instances of which are true', contrary to his view in *Experience and Theory* that 'not all valid (or identically true) forms of classical elementary logic remain valid in the three-valued logic' (p. 41) which implies that *some* forms of the former remain valid in the latter or at least that the latter does have some identically true forms. He distinguishes Kleene's logic by the set of Kleene-equivalences, i.e. biconditional formulae 'all substitution-instances of which have the same truth-value on both sides of the connective of biconditionality' on the Kleene interpretation, as I suggested. But empirical discourse seems to conform to the classical equivalence of formulae which are not Kleene-equivalent. Secondly, he has missed the formal similarity between neutrality and contingency and its bearing on the relationship between Kleene's logic and classical logic, partly by mistaking my contention regarding contingency. Thirdly, his distinction between neutrality and undecidability is muddled.

Although every familiar named equivalence of classical logic is a Kleene-equivalence, there are many pairs of classically equivalent formulae which are not Kleene-equivalent, because every contingent classical truth-function of two arguments, with the exception of exclusive disjunction and biconditional, is associated with four Kleene-functions by allowing an alternative between the non-classical value and one classical value for two out of the five non-classical sets of argument-values. (Since the value of every formula containing only one variable is identical in the only non-classical row, all classically equivalent formulae containing one variable are Kleene-equivalent). It is interesting that tautology and contradiction are associated with sixteen Kleene-functions each, in accordance with the principle that a Kleene-function has a classical value w for a mixed pair of argument-values containing classical value v only if it is assigned w for both the classical pairs having v in that position. Thus the Kleene-table for $(p \vee \sim p) \cdot (q \vee \sim q)$ and for $(p \cdot \sim p) \vee (q \cdot \sim q)$ has 'X' in all the four non-classical rows which allow choice, whereas the table for $p \vee \sim p \vee q \vee \sim q$ and for $p \cdot \sim p \cdot q \cdot \sim q$ has no 'X' in them. Even $p \vee \sim p$ and $q \vee \sim q$ are not Kleene-equivalent. A formula differing in the number and positions of 'X's in any permitted manner from a given classically tautologous formula may be obtained from the latter either by 'disjoining' it with one or more of the variables and their 'negations' or else by 'conjoining' it with some other classically tautologous formula. A classically contingent formula is, therefore, not Kleene-equivalent to its 'conjunctions' with some classically tautologous formulae and to its 'disjunctions' with some classically contradictory formulae. Similarly, $p \vee [(p \supset q) \supset q]$ is classically equivalent but not Kleene-equivalent to $(p \supset q) \supset q$ which is not Kleene-equivalent to $p \vee q$ but to the latter's 'conjunction' with $q \vee \sim q$. (My remark on this point on page 214 in my article is misleading.)

The absence of Kleene-equivalence between classically equivalent formulae considered above would, however, refute rather than support Körner's claim that empirical discourse employing inexact predicates follows Kleene's tables. Körner gives no reason for denying tautologous character to every statement containing inexact predicates, i.e. truth irrespective of whether or not any or all the atomic components are neutral. And there seems no reason at all for denying to ' x is red or x is not red' and ' x is a tree or x is not a tree' even the sort of equivalence Kleene's tables assure for 'It is not the case that x is a tree and x is red' and 'Either x is not a tree or x is not red'. His view that neutrality is essentially provisional favours the customary view that the substitution of atomic statements containing inexact predicates and the familiar logical words for the variables and logical symbols in a classically tautologous formula always yields a true statement, since such a statement remains true, whatever the combination of classical truth-values for its neutral atomic components. If every statement containing inexact predicates and having the 'structure' of a classical tautologous schema is true, then the 'conjunction' of such a statement with any contingent empirical statement will have the value of the contingent 'conjunct'. It cannot be maintained without argument that the identity of value in this case cannot be carried over to empirical discourse while that between Kleene-equivalent classically contingent formulae, whose 'biconditionals' are basic laws of classical logic like De Morgan's, Transposition, Exportation, is carried

over. If the identity of value between a classically contingent formula and its 'conjunction' with a classically tautologous schema is admitted for every set of values of atomic components containing inexact predicates, it cannot be denied to any pair of classically equivalent contingent formulae.

Körner's statement that 'the three-valued logic is itself an idealisation' (*R*) may give the impression that he regards the absence of Kleene-equivalence between so many classically tautologous formulae as a short-coming of this logic which makes it an idealisation of the relationships between statements containing empirical components. But in his book it is 'the inexhaustible and ever-changing fullness and complexity of any natural language' (p. 232) rather than the inescapable inexactness of *empirical* predicates which, he says, makes Kleene's logic 'a simplification and idealisation rather than a faithful exhibition of empirical discourse'. In any case, the above mentioned shortcomings of Kleene's logic are irremediable only so long as a simple truth table logic is sought, which must assign the same value to a schema pOq for the same set of values of the components without discriminating between a component's value's dependence and independence of the value of the other components.

Once a systematisation of the relations between the 'truth values' of statements containing inexact predicates is settled, its relationship to classical logic is not difficult to determine. The two cannot be identical in so far as no symbol of the classical propositional calculus can be interpreted as the property of being neutral. But from the fact that classical logic is not the instrument for deducing statements from those ascribing neutrality, etc., to statements, it does not follow that it is not an instrument or even the right instrument for deducing statements from statements which are themselves neutral. Nor would such a conclusion follow from the supposition that Kleene's logic or some modification of it does the first job. In the sense in which classical logic cannot be said to do this job, it does not do it for statements ascribing classical contingency, etc., either; and Kleene's tables are satisfied by the contingency-triad exactly in the cases in which they are satisfied by the neutrality-triad, as I said in the last sentence of my article, if the absence of Kleene-equivalence in the above-mentioned cases is a shortcoming. But Körner finds no difficulty in accepting 'classical logic . . . as an instrument for the deduction of contingent conclusions from contingent premises' (p. 47). The analogy with the well-known relationship between classical logic and the logic of statements ascribing classical contingency, etc., shows that it is naive to suppose that 'in classical logic all propositions . . . are assumed to be true or false' (*R*) so as to exclude even provisional neutrality, and to infer that the question of deduction from a neutral statement in accordance with classical logic 'does not even arise' (*R*) and that 'whenever inexact concepts are accommodated in a theory (embedded in classical logic) they are *ipso facto* modified by the removal of their inexactness' (*R*), unless Körner's 'neutrality (in my sense)' (*R*) is concerned with something unfamiliar rather than with the limits of sense-acuity. Whether classical logic is an instrument, and the right instrument, for deducing conclusions from statements containing inexact predicates is decided by considering whether the apparatus of classical logic can cover all possible structures of statements containing such predicates and whether it is truth-preserving and simpler than any alternative logic, and not whether these statements admit of a tripartite division.

As I explained in my paper, Kleene's tables are so related to classical logic that they would not apply even in the cases they do if classical logic did not apply to statements containing inexact predicates. Kleene's tables show that the inference of a 'conjunct' from its conjunction is truth-preserving and that if q can be inferred from p but not from any conjunct of a statement 'equivalent' to p , then q is neutral if p is. Körner mistakes my contention by missing the expression 'if the essential premises . . .' which immediately precedes the statement he quotes. Since for every set of premises a unit set consisting of their conjunction can, ordinarily, be substituted, the distinction between essential and inessential premises can only be made by considering the conjuncts as distinct premises. The validity of the inference of 'the non-contingent conjunct' (R) from its conjunction with a 'contingent' statement, therefore, does not show me to be 'plainly mistaken'.

As a 'method among others' (p. 40) the form-independent method 'of defining and determining logical truth in the three-valued logic' (p. 40) given by Körner cannot indicate an awareness of the total absence of identically true forms in this logic. And the notion of a *posteriori* logical truth it defines by confounding in effect the relation between a statement and its 'non-neutral residue' with the part-whole relation can hardly be saved by characterising it as 'a weak notion of logical truth' (R).

Körner's statement that an undecidable candidate for the membership of a class may be made decidable by supplementing the qualifications and disqualifications (p. 28) suggested that he might regard 'loosely defined' predicates of ordinary discourse like 'motor car' as admitting undecidable candidates. But he seems to apply this predicate only to formulae or statements in relation to classes of theorems. But given that 'in an axiomatic system, say A , neither p_0 nor $\neg p_0$ is deducible' (p. 93), it is wrong to say of p_0 that 'neither the admission nor the refusal accords with the rules' (R), because in that case both are refused admission. Deductive undecidability has no analogue in the realm of empirical existents, however. But to say that an object is 'neutral if, and only if both admission and refusal accord with the rules' (R) is to allow objects to be neutral candidates for the application of a predicate introduced by an inconsistent set of rules. And it does not assure that the qualifications and disqualifications for the application of a predicate having or admitting of neutral candidates are complete in the sense of being incapable of supplementation by ostensive or verbal means and that the predicate is good enough for the scientific experimenter. To say that the rules need only be 'complete in the sense that every candidate for membership is positive, negative or neutral (but not undecidable)' (R) is hardly to give any sense to 'complete', at least while the meaning of 'undecidable' is itself not distinct. And Körner does not explain the difference between the rules permitting neutral candidates and 'rules for dealing with neutral candidates' (R), which might indicate how he distinguishes between the inexactness of *empirical* predicates and the inexactness of the predicates confined to *ordinary* discourse.

DHARMENDRA KUMAR

New Delhi

GÖDEL'S THEOREM IS A RED HERRING

Lucas (1961), in an interesting and provoking article, has stimulated much discussion amongst logicians, but I still do not believe that his conclusions are correct.

I have argued (Good, 1967) that the assertion that *human logic can do some things that a Turing machine cannot do* cannot be proved by means of Gödel's theorem. Lucas (1967) misrepresents me when he says that I 'deny that there could be any peculiarly mental powers, since if there were, they could be described, and if they could be described, a computer could be programmed to simulate them'. What I maintained was merely that 'mentalism' *does not follow from Gödel's theorem*. I argued further that the essence of the matter is contained in transfinite counting and I think it is misleading to put the emphasis on Gödel's theorem. It is also necessary to point out that the human cannot know that the Gödelian formula is true: he can at best believe it by believing that the formal system he starts with is ω -consistent. (A system is said to be ω -consistent if there is no sequence of propositions $P(1), P(2), \dots$, each of which is provable in the system, and such that $(\exists n) \overline{P(n)}$ is also provable in the system. Sometimes I and others have written 'consistent' when we should have written ' ω -consistent'.)

In a way, the Gödel propositions are irrelevant to the argument since all we need to append at each stage is 'the formal system so far defined is ω -consistent'. We know from Gödel's work that this statement will not itself be provable in the system defined so far if it is true. As far as the argument is concerned the Gödel propositions can be replaced by these assertions of ω -consistency, and the Gödel construction C can be replaced by the simple process of appending these assertions. This makes it clear that everything really boils down to an analysis of the process of counting, if questions of 'belief' are left on one side.

The need to bring in *transfinite* counting is made clear by the fact that a machine could, by using the Gödel construction C , together with a form of 'triangularisation', print the sequence of Gödel propositions $G_1, G_2, G_3, \dots$ without ever stopping, so there is no question that Lucas's original argument must be taken beyond G_ω . I think his second note (1967) admits this so there is no further dispute on this point.

Everything now hinges on whether the 'creative' process of transfinite counting can be used to demonstrate 'mentalism'. Any given programme for transfinite counting can 'in principle' be improved in the sense that *there exists* another programme that can go further and it is on this basis that a mentalist should build his case. For this purpose Gödel's theorem is a red herring, and Lucas's first article should have been called 'Minds, machines and transfinite counting'. What he must prove is that he personally can always make the improvement: it is not sufficient to *believe* it since belief is a matter of probability and Turing machines are not supposed to be capable of probability judgments. But no such proof is possible since, if it were given, it could be used for the design of a machine that could always do the improving. This is impossible since it would lead to the conclusion that τ , the smallest non-constructible ordinal, is constructible. The ordinals less than τ are enumerable but not constructibly enumerable. (I am indebted to Dr David M. R. Park for this comment.) This gives yet

another transformation of the problem: the mentalist must show that he can constructibly enumerate non-constructibly enumerable sequences! The emulation of ZEUS_r is an impossible task.

I. J. GOOD

Virginia Polytechnic Institute

REFERENCES

- LUCAS J.R. (1961), Minds, Machines and Gödel. *Philosophy* 36, 112-27.
 GOOD I.J. (1967) Human and Machine Logic. *Br. J. Phil. Sci.* 18, 145-6.
 LUCAS J.R. (1967) Human and Machine Logic: a Rejoinder. *Br J. Phil. Sci.* 19, 155-6.

Reviews

WEYER, E. M. and FISHER, R., Eds. (1967). *Interdisciplinary perspectives of time. Ann. N.Y. Acad. Sci.* 128, 367-915.

What I shall try to do in this review is twofold: firstly, to bring the reader up to date (i.e. up to 1966) on discussions and experiments conducted on all aspects of Time; and secondly, to give a fair description of the contents of twenty-seven papers on this subject which were read in January 1966 at a Meeting of the New York Academy of Sciences, at a conference devoted entirely to the Problem of Time. I shall probably fail in both attempts.

Since the majority of the participants at the Conference were biologists or psychiatrists, more than half the papers (viz. 16) pertain to these subjects, while eight papers discuss the time of physics and cosmology and have more direct philosophical import. I shall begin with a discussion of the latter.

I

The two philosophically most interesting papers on the Concept of Time are those by G. J. WHITROW ('Reflections on the Natural Philosophy of Time') and by A. GRÜNBAUM ('The Status of Temporal Becoming'). Whitrow believes that philosophers and physicists have been divided in their attitude towards the following two beliefs:

- (1) that there is nothing in the external world corresponding to the distinction between past, present and future; these are therefore essentially *subjective*; and
- (2) that there is nothing in the external world corresponding to our sense of *time-direction* (time's 'arrow'); this is therefore due to memory and consciousness (and therefore subjective); while in physics the notion of 'symmetrical time' suffices.

Whitrow himself rejects both (1) and (2) as false. He says that some writers, like H. Mehlberg, accept both; while others, like Grünbaum (and, I suppose, Reichenbach) accept (1) and reject (2).

Regarding the second tenet, (2), Whitrow states that the time of physics, including statistical descriptions, has hitherto been taken to be symmetrical; but that very recently some experiments with K_s -mesons have shown that the C-P-T-invariance principle does not hold (C=charge, P=parity, T=time-reversal), and that therefore this process would not be symmetrical under time-reversal. The alternative explanations of the decay experiment which have been offered by some workers in this field are rejected by Whitrow as being insufficiently substantiated; he says that in the present state of knowledge, we must accept that not all the fundamental non-statistical laws of physics are symmetrical as regards time.

He is more sceptical regarding other evidence offered for time's 'arrow': entropy does not suffice to provide us with it (unless we superimpose another 'mark' into the process, as Reichenbach has shown); nor does Fong's theory of

spontaneous processes; nor the expansion of the universe with its redshift; nor Gold's concept of the universe as a non-reflecting sink for radiation. All these, I agree with Whitrow, may only be phases within a larger oscillatory process of the universe as a whole, although I find it difficult to see why time should be reversed at the point of return of the pendulum, or what such reversal would be like. Whitrow proposes that, instead, the direction of time may be gathered from, and dependent upon, the universe's *initial conditions*—an interesting thought which, however, he does not further develop here.

So far, so very good. But the inability to account for a unilateral direction of time from physical evidence now prompts him to seek it in the earlier-later relationship between events, which he somehow co-ordinates with the 'distinction between past, present, and future (that) lies at the core of the concept of time'. Having said at the outset of his paper that we must distinguish between (a) the anisotropy of time (i.e. time's irreversibility) and (b) 'becoming', i.e. the happening of events, and *hence** the distinction between past, present, and future, he now turns to (b), the concept of 'becoming', as evidence for (a), time's 'arrow'. He says: 'Once we abandon the distinction between past, present, and future, there would seem to be little reason for believing in the objectivity of time's arrow.'

Whitrow thus believes that the 'now' is an objective moment of becoming, where something turns from indefiniteness to being determined, i.e. definite. The future is the realm of possibilities, but actualities are only experienced in the present: the 'now' has objective reality. He quotes Hugo Bergmann's criticism of Reichenbach (in 1929): 'The "now" is private to each observer. "Determined" means "determined for me now".' In reply, Whitrow devises an example to show that the 'now' can be shared, can be made intersubjective, even objectivised—by using a camera as the second observer. But, clearly, this only shows that the camera and I have the same 'now'. Since the camera is my percept, the picture it takes will of course be like my own perception, and will refer to the same moment of time. And if it isn't, if nobody can verify at which moment the picture was taken—do we then know which 'now' it is that is objectivised? Does the word 'now' then still have a reference?

It is true that there is a moment of becoming, where something that was not the case turns into being the case—e.g. a piece of litmus paper inserted in acid will turn red, and be red thereafter. But this moment of 'becoming' can lie in the past, it can be the present, or may occur in the future: it does not provide us with a criterion for distinguishing between the three temporal regions. Nor can it be identified with a universal 'now', for the same reason: that there have been 'nows' in the past, as there will be in the future, since the term is always speaker-reflexive. The third concept, that of time's 'arrow', can certainly not be derived from either 'becoming' or the 'now', even if this were intersubjective; its objectivity could only be derived from empirical evidence, not from metaphysical considerations.

While Whitrow thus makes a strong plea for the objectivity of 'becoming', GRÜNBAUM takes the opposite view. He defines 'becoming' subjectively, as a mind-dependent process of experiencing an event, while being aware of this fact of experiencing. He thus distinguishes two different, though simultaneous,

**My italics.*

mental processes, one having empirical (perceptual) content, the other having 'awareness that one is experiencing something' as content. In his words: '*M* experiences the event at *t* such that at time *t*, *M* is aware of having that certain experience *simultaneously* with being aware of the fact of having it at all.' The immediate consequence of this is, in his own words, that 'it makes *nowness* (and thereby *pastness* and *futurity*) depend on the experience of conceptualized awareness that an experience is being had . . . '.

In simpler words: the becoming of the 'now' is *made* mind-dependent by using the word 'awareness' in its definition. Grünbaum is explicitly *not* concerned with the becoming of physical events, e.g. the coming-into-existence of a child from non-existence (he calls that 'tenselessly beginning'), or of an explosion of gunpowder, when ignited. (These examples are his.)

Likewise, indeterminacy, for him, is no proof of coming-into-being of what was not there before, because while it was indetermined, the future of the particle was *determinate*, since the indeterminacy is still relational, not absolute.

He seems to jump from the true belief that 'past' and 'future' are subjective, speaker-reflexive terms, to the 'fact' that there can be no independent (objective) physical occurrence called 'becoming', since by definition this involves consciousness. He defends himself against the criticism that according to his theory of 'becoming', physical events would be mind-dependent. No, he says, only the *awareness* of physical events counts as 'becoming'—namely, a becoming aware of their happening now, i.e. in the observer's present. This, I find, is both, trivial and false: trivial insofar as awareness, like perception, always takes place in the observer's present—his 'now'—and no conclusion about physical becoming can be drawn from it; and false, as regards our normal use of the word 'becoming' (if it has a normal use at all!), by which we usually do *not* mean a 'becoming-aware-of', but an objective coming-into-being of an event or a thing. The word is not usually taken to be observer-reflexive. Grünbaum therefore does not say in his paper whether he believes that the becoming of physical events is objective or not—he is mute on this point.

Another consequence of his definition of the 'now' is that when nobody experiences it, there is no 'now' (*vide* Berkeley!); that is to say, the time-series of his 'nows' is not dense, and one cannot derive a continuous succession of time units from it.

The third paper in this section, by COSTA DE BEAUREGARD ('Two Principles of the Science of Time'), is even more extreme than Grünbaum's on this point. He holds a completely deterministic and static view about time. All events are 'there', we only come across them in time. He believes in the so-called 'block universe' in Minkowski's space-time. 'The space-time continuum of Special Relativity is an unchangeable four-dimensional block', and 'the world's history is written once for all *sub specia aeternitatis* in divine thought'. This view of Poirier is, he thinks, further borne out by the relativistic formulation of Quantum Mechanics. According to Feynman, creation and annihilation of electron-positron pairs are events written once for all in space-time; 'a good integral over all space-time implies a fatalistic philosophy'. Beauregard seems also to agree with Schrödinger that Heisenberg's Indeterminacy is only epistemic, not ontological, indeterminateness. While the hitch-up of quantum-mechanical experiments *guarantees* that one quantity becomes unmeasurable

when the other is being measured, 'the result of a quantum-mechanical measurement is an *objective event* which is written in space-time *sub specia aeternitatis*'.

He agrees with Grünbaum that probability and fatalism are compatible in QM, and that the probabilistic character of QM has been reconciled with Minkowski's four-dimensional space-time geometry; free will is possible even in the 'block universe'. Past, present, and future are therefore totally subjective. He (also) quotes Hugo Bergmann: "'Now" is the mode of the experiencing ego.'

Beauregard makes the interesting and important point that entropy increase and wave retardation stand in a one-to-one correlation, particularly for quantised waves, since scattering (of waves *and* particles) is a way of increasing entropy. Physical irreversibility is thus fact-like, not law-like (as Mehlberg has observed).

He draws attention to the close connection between information and probability: every gain in information is paid for by a loss of entropy. In an isolated system, the sum of neg-entropy and information cannot go up. Life implies an incoming flux of information and thus an anti-entropic tendency. Since observation is 'easier' than action, causality is more obvious than finality and prevails over it—i.e. the entropy curve goes up.

Apart from the last bit, which is open to criticism too obvious to be worth making, the paper is highly informative and betrays great knowledge of the subject; but it is written more as a presentation of prevalent views on time than as a discussion of a thesis of his own. What are the Two Principles of the Science of Time? I don't know; he doesn't say.

In contradistinction to Beauregard's mundane, but unspecific article, the last of the four papers dealing with the Concept of Time is very specific, highly condensed, and quite original. GOTTHARDT GÜNTHER's (Illinois) 'Time, Timeless Logic, and Self-Referential Systems' approaches the dichotomy between subjective and objective time from a logical point of view. Since Parmenides and Aristotle, philosophers have tried to divide the world into Thought and Being; Time falls between these two chairs, and there has therefore always been a tendency to eliminate it, or to objectivise it by spatialising it.

What's wrong here, Günther thinks, is the logic: two-valued, with a *tertium non datur* between assertion and negation, it cannot account for the temporal aspects of becoming, which disobey the exclusor. He therefore suggests a multi-valued logic to encompass Time. It avoids the dilemma that faces the category of 'pseudo-objects' which imply their contrary: relations like right-left, before-after, positive-negative, and concepts like life (death), motion (rest), and past (future)—as against 'ortho-objects' which exist (have meaning?) independently of anything else. Hitherto, only the future has been held to be indeterminate, i.e. three-valued, by some (eg. by Reichenbach), whereas past and present were believed to obey the 'law' of the Excluded Middle.

Günther introduces the notion of a 'Kenogram', a concept something like the inverse of a logical function (e.g. if $y=f(x)$, then the inverse is function $x=\phi(y)$). The two symbols he uses are an empty and a filled triangle (Δ and ∇) for the two functions, respectively. He then sets up tables of these kenograms, where more than two variables yield a matrix of n^n values (for $n>2$), some of which may be represented by empty places.

In this way, time can be assigned an ontological value of its own—'Becoming'. The difference between Being and Becoming, he says, is that between deutero-

grammatic and morpho-grammatic structure which lies at the root of Zeno's troubles, who tried to reconcile Change and Motion with the permanence of Being and got a paradox. This can be avoided by a multi-valued logic. We also need it to represent Being, Thought, and Time. One then needs a further locus for the 'subject of cognizance', i.e. the observer, who must be able to distinguish between (a) the world given in its three-valued ontology, (b) his thought-image of the world, and (c) himself as producer of this image. One therefore needs a five-valued logic for this representation. 'The personal ego constitutes itself in the triadic relation between environment, image, and image-making. The fifth locus is that of self-reference'.

Günther's paper is short and very condensed, so much so that he was asked to enlarge on it in the Panel Discussion which followed at the end of the Conference. This his colleague (and collaborator?) HEINZ VON FOERSTER did, thus producing another sixteen pages of print. It is an attempt to resolve the puzzles about time by changing the logic, and, furthermore, of applying this logic to metaphysical concepts, which is more than most other producers of new logics have done. Is it too much?

The only purely mathematical paper is by C. MUSES (Lausanne) on 'Time, Experience and Dimensionality: An Introduction to Higher Kinds of Number'; it deals with higher dimensions ($n > 8$) which are represented graphically by their diagonals. Muses introduces a different kind of dimensionality, where exponentials of imaginary numbers play a part. The paper is so highly number-theoretical that I can't begin to understand it, and shall resign myself to mentioning only three of the interesting things he has to say about time.

(1) 'The direction of time is physically characterised by a constant *minimising* of entropy-increase.' This, he says, lies behind Hamilton's Principle of Least Action. This suggestion seems to me to be the most elegant—and plausible—way of deriving time's 'arrow' from observations of entropy alone: by measuring the *rate* of entropy-increase with increasing time, and finding it to be decreasing! A most interesting result!

(2) 'There is no time except the NOW'—a contention which he deduces from the fact that in a certain complicated equation, t becomes $=0$.

(3) Even if we reverse entropy, or have an entropy-less system, 'the reversal occupies a different place in the time sequence than the process whose reversal it is'. In other words, time's direction cannot be equated to entropy-increase. This conclusion already follows from what he said under (1), and it is intuitively very acceptable; the only trouble with it is, that while such entropy reversal is not logically inconceivable, the laws of physics would be so drastically changed if it did obtain, that it seems questionable whether our mathematical definition of time, which enters into all differential equations of physics, might not also have to undergo such radical change that it would become unrecognisable.

Of the two remaining essays on Time in Physics, S. GERATEWOHL's (Washington) deals with 'Navigation in Space-Time'. He applies Relativity physics to space-flight and discusses some of the fairly well-known paradoxical consequences of such an accelerated travel. For instance, a space-craft accelerated to 1 G ($=32\text{ ft/sec}^2$) would cover 1000 earth years in 11 years of travel; after 22 years of their own time, the astronauts would return from their distant star to earth, where 2000 years had elapsed since their departure. On their return flight,

they would have witnessed the history of 1000 years on earth *yet to come!* (How such 'witnessing' is to occur, I cannot imagine!)

A seemingly more plausible consequence would be that at a speed of $0.9c$ (where c =the speed of light), the distribution of all the stars in the direction of their flight would shrink to a cone of 30° angle of subtension, at $0.99c$ to a cone of 18° . At $0.886c$, only the stars outside a 55° cone will suffer redshift, inside the cone they will appear blue. In direction opposite the line of flight, they will dilute and fade away. Only those at exactly 55° will appear normal in spectral quality.

He quotes Eugen Saenger as saying (in a paper in 1956) that the velocity of light need not be the fastest attainable: photon propulsion could produce faster motion. In that case, if then all life processes would be equally retarded as the clocks, all known points of the universe could be reached within the lifetime of a crew of astronauts. Geratewohl comes down unambiguously this side of the clock paradox by accepting the belief that clock retardation can be meaningfully applied to biological systems as well.

R. W. FAIRBRIDGE's short, but very interesting paper on 'Geological and Cosmic Cycles' is a compendium showing correlations between sunspot-cycles (of 11, 22 and 44 years) and ultra-violet, X-ray, and other radiation 'crescendos' which strengthen the earth's thermal blanket (by producing more ozone from oxygen), thereby raising the earth's temperature by 1°C in fifty years—the well-known 'greenhouse effect'. The mean height of the sunspot peaks has *doubled* in the past fifty years. The sunspot variation also forms a larger (Gleissberg) cycle of roughly eighty years, which corresponds vaguely to the age-old '100 year calendar' periodicity of the earth's climate.

There is a correspondence between recent establishment of the galactic rotation of 216,172,800 years with the geological cycle of about 200,000,000 years which has been established for some time. Another important co-ordination is that between the rate of C^{14} formation which varies periodically (by about 2 per cent over a century or two), but inversely, with sunspot maxima.

Before turning to the biological papers, one article by HERMAN WOLD (Uppsala) should be mentioned: 'Time as the Realm of Forecasting'—an impressive, though philosophically unyielding paper on forecasting models. It has little to do with time as such, which he presupposes to be flowing uniformly, except that it deals only with the future.

He divides forecasting models into four main categories:

(1) the descriptive experimental; (2) the explanatory experimental; (3) the descriptive non-experimental, and (4) the explanatory non-experimental.

Descriptive approaches attempt an answer to the question 'What?', and are defined as univariate; while explanatory models answer the question 'Why?' by giving cause-and-effect relationships.

Examples for (1) are two-valued, and the best known: e.g. Bernoulli's urn scheme, or prediction of the sex of an unborn baby; under (2) he mentions Jos. Dietl's report on pneumonia, biological prediction of the basis of Mendel's laws, and betting examples with dice. Among the group (3) falls a Gallup poll before a presidential election, and among (4) are the so-called 'Engel curves' and income (if income rises, the consumption of goods goes up), or the relation between sunspot frequency and radio storms.

The majority of papers read at the conference deal with perspectives of time in the life sciences. They fall into two groups: six of them discuss biological processes, such as ageing, or diurnal rhythms in living organisms; while ten papers deal with the psychological awareness of the passage of time and the way this time-sense is affected by psychosomatic drugs. Unfortunately, space does not permit the discussion of more than five of these psychological essays.

A. Four of them, viz. those by H. LEHMANN, S. GOLDSTONE, R. EFRON, and R. FISHER, centre about psychological experiments on the estimation of very short time intervals—mostly between 0.15 and 4 seconds—conducted by Goldstone with normal and psychopathic individuals. I shall therefore discuss these together, beginning with GOLDSTONE's paper 'The Human Clock: A Framework for the Study of Healthy and Deviant Time Perception'.

Man's basic time perception is bipolar, according to Goldstone: he perceives and measures time as long/short, before/after, more-or-less. The 'human clock' is therefore relative, not absolute, has no fixed zero-point, and undergoes changes according to different 'levels of adaptation'.

In his experiments, very short visual and auditory stimuli were to be judged as more, less, or equal to one clock second. All stimuli were anchored either by short light or sound 'blimps'. The results showed that on the average, auditory stimuli were judged to be considerably shorter than visual stimuli of the same duration. And the most perfect time estimation was achieved when one short auditory stimulus was given. But when experimental conditions were changed, e.g. a different series of stimuli were given, the results differed considerably.

The conclusions Goldstone draws from his experiments are the following: (1) The time we hear seems longer than the time we see—i.e. we tend to underestimate the length of light flashes; and (2) our internal reference scale is shifted with great ease in correspondence to experimental conditions. In other words, our internal clock is not absolute, but highly relative.

This fact does not seem to be much changed in psychosis: the acute schizophrenic, according to Goldstone, tends to overestimate short auditory stimuli, but not the visual; while the chronic schizophrenic shows little aberration in time estimation from the normal, and the difference between auditory and visual stimuli is not affected by schizophrenia.

Young (normal) children tend to overestimate very short durations, while they underestimate longer ones (all within the range of 0.15 to 4 seconds). Experiments thus point to a primacy of auditory learning in earliest childhood and a predominance of vision in adult perception.

HEINZ LEHMANN ('Time and Psychopathology') says in addition that hysterical patients and psychopaths give longer time estimates than normal, while duration is underestimated by neurotic and psychotic depressives. He also finds that if a person has been trained to overestimate (or underestimate) time consistently, his motor inhibition becomes affected. Conversely, experiments with rats have shown that if lesions affecting motivation have been produced, the animals' time behaviour is affected.

In his article, Lehmann distinguishes four types of time experience:

- (1) External time calculation; (2) Internal time estimate;

(3) Internal time awareness; and (4) Internal time perspective.

(1) *Calculation of time* is the learned reading of an instrument (clock, calendar) and is impaired only under the most severe mental disorders.

(2) *Time estimation*—which was tested in Goldstone's experiments—develops between the ages of 5 and 9 years and remains fairly stable until the 7th decade of life.

In amplification of Goldstone's results, Lehmann says that the minimum interval at which we can judge *visual* stimuli as successive is 0.02 to 0.05 seconds; below that, they appear as one continuous stimulus. Yet, an interval of only 0.002 seconds between two *auditory* stimuli still permits us to judge them as separate.

On the other hand, G. SCHALTENBRAND ('Consciousness and Time') states that in order to have permanence of visual perception, the minimum duration was found to be 0.11 seconds. A maximum interval of 70 milliseconds between pictures will let us perceive a film as continuous—above that, it begins to flicker. The shortest time duration in which we can take in a piece of information (e.g. a word) is 1/10 of a second; and we can absorb 4.5 to 6 'bits' of information per second before the central nervous system becomes over-loaded.

ROBERT EFRON says on this point in his paper 'The Duration of the Present' that all perceptions shorter than 66 milliseconds seem to us to have no duration at all, but appear as instantaneous. His crucial experiment is one where a flash of red light is followed by a green flash, and the resulting perception is one of *continuous* yellow, if the flashes follow each other at an interval of less than 20 milliseconds! At larger intervals, the true colours appear intermittently.

These results, taken together, seem to point to the 'fact' that temporal perception is quantised, and its units are recognisable and even measurable, but that they are also variable and differ within a certain range. Possibly, according to Schaltenbrand, they are themselves not elementary, but are all multiples of a tiny basic unit of time, which might, with Heisenberg, be identified with the duration of one electron spin, the shortest periodic occurrence in the world, viz. 10^{-26} seconds. But this suggestion seems highly conjectural.

(3) *Time awareness*, says Lehmann, is the impression that time moves rapidly or slowly. He cites a few experimental results: He tested seventy-six psychiatric hospital patients, divided into six groups: old schizophrenics, young schizophrenics, depressed patients, hypomaniacs, and patients with organic brain syndromes, (also old and young separately); and twenty normal people as controls. Eighty per cent of the *normal* subjects tested felt that time passed quickly, the others felt it passed at medium speed. Of the neurotic or psychotic depressed patients, 60 per cent said that it passed slowly! This is a very typical result of depression, substantiated by many tests.

When asked to estimate the duration of half an hour, more than 50 per cent of all patients, except the old schizophrenics, *under-estimated* the time. Of the old schizos, 90 per cent *over-estimated* or were correct. Euphorizing drugs, like alcohol, LSD and amphetamines, produce a feeling that time moves rapidly.

ROLAND FISHER in his long, comprehensive article (for which he gives over 200 references) on 'The Biological Fabric of Time' offers more evidence on this point. He distinguishes *data content* from *data processing*. Excitatory drugs, like LSD and psilocybine, further an increase in data content *and* data processing, and result in an experience of 'time contraction', in which what is experienced

in a short span of time *seems* to take days and years (e.g. in the few seconds during a fall from a cliff, a survivor will have 'seen' his whole life reel off before his inner eye); while, under mescaline, increase in data content *without* increase in processing is experienced as a 'torrential flood of inner sensation'. Under the influence of these drugs, metabolism (i.e. heart rate, temperature, etc.) goes *up*.

We meet the notion of 'information processing' again in Efron's important investigation of the duration of the 'now' as experienced. He says, the 'now' does not begin with the stimulus, but with our *awareness* of it. The onset of perception is introspectively given, but can often not be established, for two reasons: there is a time-lag due to the travel of the stimulus along the afferent nerves to the brain; but also, there is a time of *processing* of the information received by the brain, which takes about 60-70 milliseconds. (This figure is an average from different investigators who are in close agreement.) This throws an interesting light on his aforementioned experiment which showed that perception shorter than 66 milliseconds seems to have no duration. We have obviously no awareness of the 'real' onset of a perception. If another 'bit' of information is offered while processing is still going on, we either contract both into one perception (yellow light!) or we are blocked. He equates the duration of a perception to the experienced duration of the present.

(4) *Internal time perspective*, according to Lehmann, is a person's orientation toward past, present, and future. When his psychopathic patients were asked which was farther away for them, yesterday or tomorrow, 70 per cent of the old schizophrenics felt that tomorrow was farther away; 50 per cent of the depressives felt this, and only 21 per cent of the normal group. Here, however, the fact of the person's age seems to carry more weight than his illness: the older a man is, the more his life is oriented toward the past, and the less 'real', i.e. the farther away, the future (tomorrow) will seem. Compulsive neurotics, and depressives, are oriented toward the past for a different reason: they are blocked towards the future, because they are blocked towards action—they are afraid of it. All neurotics seem to have a double motivation towards past and future, and neglect the present, while lobotomised patients are restricted to it. It seems that old people tend to persevere in their daily routine in order to safeguard themselves against their imminent loss of temporal perspective.

Lehmann's explanation of our sense of time is very interesting: In the psyche, he thinks, entropy is constantly decreasing; this is what integration of incoming information amounts to. The randomness of our memory-store is constantly decreased by ordering. Therefore, the differential between a person's neg-entropy and his information-input is constantly diminished during life, since input rises steeply from birth and gradually decreases in old age. Thus the store of organised experience grows faster than information-input, and entropy decreases—until death, when it rises sharply. The size of this differential is the code of our memory in time. If information-input is larger than integration (processing), as it is in psychosis, we get the disruption of personality functions and of time sense and memory.

B. Turning from the psychological to the biological papers, we are also turning from the discussion of very short time intervals to the long periods of biological changes: three of the six biological papers deal with the process of aging.

BERNARD STREHLER ('Cellular Aging') investigates the bio-chemical mechanisms of this process. Having shown a reduction of various biological functions with increasing age, such as metabolic rate, maximum breathing capacity, cardiac index, filtration rate of the kidney, and the amount of cell water present—some of which decrease sharply after the age of 45—and the statistical probability of death as a function of age, Strehler looks for a genetic theory of aging. In some species, he says, as for instance in the salmon, the high mortality of the fish after spawning is advantageous to the offspring—they feed on the decay minerals of their dead parents—and thus tends to be of selective advantage to the species. But for most other species, this is not so.

Cell division decreases with age. Yet some cells, like those of the human skin, replenish themselves at a rate far beyond necessity or advantage to the organism or the species.

Some genetic effects of age do not influence maturation and reproduction, and are *therefore* (sic!) not counteracted by nature: e.g. the pigmentation in the cells (especially in the cardiac cells) of old people. There seems to be little causal symmetry between this process and aging; i.e. while increased pigmentation storage is an *effect* observed in old age, its influence on age has not (yet) been established.

Ability to mutate is a universal feature of all living forms and furthers the emergence of new species, as well as variability within the old. But the loss of the function of mutation is only 1 per cent per year during normal aging and is therefore not enough to account for it. The amount of abnormal chromosomes produced by mice under laboratory exposure to an ionizing radiation of 33 roentgen is 5 per cent—yet they can survive exposure to 500 roentgen; so that exposure to radiation (e.g. cosmic gamma rays) cannot be held responsible for aging.

Therefore, Strehler concludes, no successful genetic theory of aging has yet been produced which could account for a process that is—or should be—the immediate concern of all men alive.

A less teleological approach and a more optimistic view is presented by GEORGE SACHER ('The Temperature Dependence of Development and Aging'). He believes that aging can be shown to be due to, or at least corresponding to, an increase of organisational entropy production as *temperature* is displaced from an optimal value. His investigation is restricted to insects, and in particular to the *Drosophila* fly, whose optimal environmental temperature is 25°–30°C for maturation of egg, pupa, and larva, and for metabolic rate. Above and below this temperature range, mortality rates go up sharply and entropy production, per calorie, increases. Other processes having a definite temperature optimum are mutation, developmental anomalies and total developmental failure (obviously, all these are least at the optimum). Sacher concludes that 'aging is no longer a question of *how much* metabolic work, but of *how well* this work is done, in thermodynamical and informational terms'.

His paper is difficult to understand, and the graphs he produces to explain it are on a logarithmic scale and plot inverse absolute temperature, and therefore do little to lighten one's task.

A very interesting paper with more direct bearing on temporal concepts is that by J. E. GANDER on 'Some Relationships between Steady-State Kinetics and Time'. How is it possible, Gander wants to know, for an organism to main-

tain its structure and yet obey the 2nd Law of Thermodynamics? The answer is sought in time-dependent chemical and physical interactions within the cell. There exists a system of reactions of allosteric enzymes in a metabolic sequence within an organism. A positive entropy flux obtains through interaction of non-linear oscillating concentrations of these metabolites. He suggests that the kinetic parameters of the regulating enzymes have a direct influence through their saturation and dissociation upon the energetics of living cells and thereby upon the process of aging.

He suggests several models of how this entropy increase may be effected: (a) The typical 'black box' model demonstrates how energy is fed in (in the form of nutrients) and is eliminated to a lesser degree (as wastes). Entropy is increased by catabolism of nutritive material as well as by non-linear oscillations of the components of the system, so that heat increase and entropy increase is always positive.

(b) Another model shows the coupled oscillations between enzymes in a phase-shift, where one enzyme catalyses the formation of one substrate, J' , and inhibits that of J'' , while another enzyme does the opposite. Since there is feedback between them, there results an oscillation about a steady-state which is time-independent.

(c) Thirdly, there are allosteric saturations which act as amplifiers, so that the velocity of reaction increases faster than does the substrate concentration. Therefore, the kinetic energy of the system increases, and with it its entropy.

These models predict that the rate of aging is a function of the accumulation of aberrant allosteric enzymes, resulting from a mutation of DNA coding for these proteins, and from errors in 'transcription'. If they are correct descriptions of what goes on in a cell, then it would follow that aging results from an accumulation of cells that become increasingly less efficient, leading to a partial and eventually complete failure of the organ.

Perhaps the cleverest of the biological papers, all of which are very good, is that by B. C. GOODWIN (University of Sussex) on 'Biological Control Processes and Time'. He poses the question whether biological systems (e.g. cells) are active and have *dynamical* quality, much the same as physical systems. He believes so, and proceeds to show in which way.

Cellular processes are unstable, he says, since there is a constant change or motion in the concentration of molecular species. This motion consists of a continuous oscillation about a steady-state and manifests itself in biological rhythms and periodicities.

The origin of this oscillation is the feedback activity of enzyme systems, acting on or repressing their substrates, among them the messenger-directed proteins (*vide* the foregoing paper by Gander!).

Goodwin has worked out the mathematics of this dynamic activity for the growing cell and presents it in the form of a differential equation, where the variables are the volume V_t , enzyme production X_t , or protein concentration Y_t , and the number of molecules of the species N_t . He gets a quadratic equation, which shows that the growth within the cell is exponential—as expected.

He has also investigated bacteria and found that they, too, operate under negative feedback control loops in oscillation, and that their frequency is about the same as the cell division cycle. These oscillatory control signals gradually

produce synchronisation of control processes, which give rise to large periodic changes in metabolite concentrations. Therefore, spontaneous *synchrony* develops in a bacterial culture.

This result is of great importance, as it shows *for the first time* that bacteria do *not* present an exception to the periodic phenomena in biological systems, as was hitherto believed.

The evolution of cells is, then, a result of a spontaneous selection of appropriate frequencies which operate so as to achieve a state of *maximum temporal order*. The macrocosmic manifestations of this fundamental rhythm is the cell division cycle, and its oscillatory energy is an analogue to the dynamic energy acting in physical systems.

The two other papers in this section also deal with periodic cycles in biological systems. A. SOLLBERGER ('Biological Measurements in Time') discusses several such periodic processes of which I shall mention only two.

(a) Diurnal temperature changes in the liver of young chicks are recorded. A high temperature is found to obtain during the day, a low one at night, with a sharp drop at dusk and a more gradual rise at dawn, and a maximum at 12-2 p.m.

(b) In a much more complicated experimental series, the output of potassium in the urine of healthy and paraplegic men is measured in hourly samples, as an indication of their metabolism. The subjects are first put on a 7-hour feeding period, and after 10 days' adjustment to this, on a 19-hourly period (with constant lighting, but with a clock available to them). In the beginning of each period, the former rhythm is superimposed on the potassium output, but it is lost after about 5-7 days, when the system adjusts to the abnormal hours of feeding, and the curves then vary periodically with food intake. The food stimulus thus acts as a synchronising influence.

Less technical, but at least as interesting, is the short article by ERWIN BÜNNING on 'Principles of Biological Chronometry' with which we shall terminate the discussion of biological papers.

Plants and animals show diurnal rhythms determined by the hours of daylight and darkness. This 'hereditary' rhythm will continue for several months, even when conditions are changed and the day is artificially lengthened or shortened. Flowering and leaf development in plants are further dependent upon the longer periods of spring and autumn, depending, that is, on whether short days are followed by longer ones, as in spring and early summer, or vice versa (late summer and autumn); while plants in tropical regions lack this response to the seasons. If plants are brought from the North to the tropics, they keep the length of their flowering cycles, but are no longer synchronised—flowers of the same species may bloom at different times, even on the same plant or tree!

Longer cycles, e.g. of the sea algae, depend on the moon (i.e. on the tides), whose diurnal cycle of 12.4 hours (and 24.8 hours), coupled with the solar cycle of 12 and 24 hours, produce 'beats' at 15 and 29 days—the lunar synodic period. This explanation had been tested and confirmed by an artificial shortening of the tidal cycle in the laboratory, when the beats were found to occur at shorter periods.

All the evidence presented in these biological papers seems to point to a great flexibility of adaptation of biological systems—including humans—to certain

selective pressures exerted by the environment on the organism regarding its temporal behaviour.

There can be no doubt that our biological and, ultimately, physical structures are time-dependent, and that they reflect this temporal influence to a degree not obvious in everyday observation; the very fact that this temporal standard is not fixed, but can be shifted with ease, after a relatively very short time of transition, to fit new environmental conditions of periodicity, is an unmistakable sign of the fact that time, for all living organisms, is very 'real' indeed.

What, then, is the outcome of such a Conference on Time? What is the *facit* of the work of almost thirty scientists, reflected in more than 600 pages of print?

As J. T. FRASER has said in his long and excellent summary at the end of the book, the Interdisciplinary Study of Time arose from a belief in its need, based on a *Principle of the Unity of Time* which assumes that (a) 'when specialists speak of time, they all speak of various aspects of the same entity', and (b) that 'this entity can be studied by the sciences, and can be made a meaningful subject of contemplation by the reflective mind'.

The immediate practical consequence of the conference was the foundation of an 'International Society for the Study of Time', of which Dr Whitrow was elected first President, and Dr Fraser its Secretary, which will meet regularly in some part of the world, to perpetuate and encourage research and discussions on the Problem of Time. One can hardly think of a more optimistic result of the success of such a conference than the unanimous decision of its participants to hold another one just like it in the near future (in 1969)!

It looks as though an age-old philosophical problem, on which many clever things and an incredible amount of nonsense has been said in 3,000 years, is finally going to be studied in a systematic and scientific way. One is looking forward to the results with eagerness and curiosity!

EVA CASSIRER

University of St. Andrews

PRIOR, Arthur N. (1968). *Papers on Time and Tense*. Oxford: Clarendon Press 30s. Pp. 164.

This is Professor Prior's third book on the logic of time and tenses, and in it he collects together thirteen papers, of which eight are predominantly informal, and five predominantly formal. Of the less formal papers all but one have been previously published, but of the formal papers only one has been previously published. It is impossible in a short review to do more than indicate some of the many points of interest which are raised. Topics discussed which should be of especial interest to philosophers of science include the logic of ending time, determinism, the metaphor of time flowing, relativised and metric tense-logics, and egocentric expressions. Of particular value is the last paper, 'Tense Logic for Non-Permanent Existents', in which Prior returns to the formal problems posed for tense-logic by admitting that if individuals only exist for a certain

time then certain propositions made about them (e.g. by the use of the demonstrative) will only exist for those times. A parallel problem for modal logic—arising because individuals exist only contingently—was first noticed by G. E. Moore. A great merit of Prior's work is his serious, formal consideration of this problem.

Prior's work on formal tense-logic has clearly been of outstanding importance, and the large literature on the modal systems which form tense logics testifies to their intrinsic interest. It is harder to say how valuable this work is in increasing our understanding of the concepts of past, present, future, before, after, and their cousins. The most substantial dispute is this: granted that (some) temporal language can be represented formally by means of a modal calculus with operators (it has been the case that . . . , it will be the case that . . . , etc.) forming things true or false at a time out of other things true or false at a time, must it be formalised like this, or is this the best formalisation of it? At the time of publication of *Time and Modality* it was clear that a great deal of doubt existed about this, and it is disappointing to find that none of the predominantly philosophical papers in this collection set forth or augment Prior's arguments in defence of tense-logic.

The chief rival of tense-logic for the favours of temporal discourse is a calculus quantifying over instants of time, with a primitive 'earlier than' relation. The formal relationship between such a calculus and tense-logics is excellently discussed in the paper 'Tense Logic and the Logic of Earlier and Later'. If we regard instants as logical constructions out of tensed facts and express this formally by supposing instant variables to be propositional variables, then the logic of 'earlier than' can be seen as subordinate to tense-logic. But even if instants are in some way logical constructions out of tensed facts, it is not at all clear to the reviewer that this should be formalised by treating instants as a sort of proposition. After all phenomenalism takes physical objects to be logical constructions out of sense data, but does not take a physical object to be a sort of sense datum. This is very serious, because if little philosophical weight can be attached to the manoeuvre of treating instant variables as propositional variables, then tense logic will not be adequate to formalising things which we want to say about instants of time being earlier than other instants. This is perhaps especially serious when the semantics of the logic demand understanding of the notion of being true at an instant, and this point by itself renders odd Prior's claim (Paper XII, 'Quasi Propositions and Quasi Individuals') that perhaps tense-logic gives the first order theory of instants what sense it has.

The most substantial single issue treated in the less formal papers is that of determinism, with which four of the eight are concerned. Professor Prior rightly wishes to take seriously the idea that the future is open to alternatives in a way in which the past is not, but his view that this could be achieved by supposing that future propositions which are not now inevitable are not now either true or false is not very convincing. The long paper 'The Formalities of Omniscience' in effect shows that it is inconsistent to hold all of the following three propositions: it is possible that something should know all true propositions; it is impossible that something not yet inevitable should be known now; some things not yet inevitably true are true. However, here the mixture of formalism and historical data really does obscure the philosophical points

which should be made in deciding which of these three propositions should be denied.

Nevertheless, the topics discussed are always important, and Prior's enviable formal ability seldom obscures his views. It seems more than likely that future work on temporal discourse must be based on an understanding of them.

S. BLACKBURN

Churchill College Cambridge

Recent Publications

RECENT PUBLICATIONS ON THE PHILOSOPHY OF SCIENCE

(a) BOOKS RECEIVED FOR REVIEW

- Angelelli, Ignacio (1967). *Studies on Gottlob Frege and Traditional Philosophy*. Dordrecht, Holland: D. Reidel Publishing Co. 96s. Pp. xvi+290.
- Arnold, Wilhelm & Zeltner, Hermann, Eds. (1967). *Tradition und Kritik: Festschrift für Rudolf Zocher zum 80. Geburtstag*. Stuttgart: Friedrich Frommann Verlag. DM 66. Pp. 352.
- Blokhintsev, D. I. (1968). *The Philosophy of Quantum Mechanics*. Dordrecht, Holland: D. Reidel Publishing Co. \$9.00. Pp. viii+132.
- Brooks, H. (1968). *The Government of Science*. Cambridge, Mass: M.I.T. Press. n.p. Pp. viii+343.
- Cardwell, D. S. L., Ed. (1968). *John Dalton and the Progress of Science: Papers Presented at a Conference in Manchester, September 1966*. Manchester University Press. 55s. Pp. xxii+352.
- Chattopadhyaya, Debiprasad (1967). *Individuals and Societies: A Methodological Inquiry*. Calcutta: Allied Publishers. Rs 16.00. Pp. vi+196.
- Cohen, I. Bernard (1968 reprint). *Franklin and Newton: An Inquiry into Speculative Newtonian Experimental Science and Franklin's Work in Electricity as an Example Thereof*. Cambridge, Mass.: Harvard University Press. 57s. Pp. xxvi+657.
- Cohen, Robert S. & Wartofsky, Marx W., Eds. (1968). *Boston Studies in the Philosophy of Science*, Vol. 3. New York: Humanities Press. n.p. Pp. xlix+489.
- Ehlers, H. (1968). *Logic by Way of Set Theory*. New York: Holt, Rinehart & Winston, Inc. n.p. Pp. xiii+386.
- Grünthal, Ernst (1968). *Psyche und Nervensystem*. (Erfahrung und Denken, Band 27.) Berlin: Duncker & Humblot. DM 13.60. Pp. 88.
- Helmholtz, Hermann L. F. (1968). *Helmholtz on Perception: Its Physiology and Development*. Ed. Richard M. Warren & Roslyn P. Warren. London: John Wiley & Son. 80s. Pp. x+277.
- Humphreys, Willard C. (1968). *Anomalies and Scientific Theories*. New York: Freeman, Cooper & Co. \$5.00. Pp. 318.
- Jolley, J. L. (1968). *Data Study*. London: Weidenfeld & Nicolson. Paperback 14s. Hardback 30s.
- Klein, J. (1968). *Greek Mathematical Thought and the Origin of Algebra*. Cambridge, Mass.: M.I.T. Press. 117s. Pp. xvi+360.
- Knox, Malcolm (1968). *Action*. London: George Allen & Unwin. 45s. Pp. 250.
- Lazerowitz, Morris (1968). *Philosophy and Illusion*. London: George Allen & Unwin. 48s. Pp. 262.
- de Mauro, Tullio (1967). *Ludwig Wittgenstein: His Place in the Development of Semantics*. Dordrecht, Holland: D. Reidel Publishing Co. 36s. Pp. viii+62.
- Mayer, Charles Leopold (1968). *Man Faces His Destiny*. Trans. Heloise Norwell, J. S. Norwell & D. C. Fox. London: Johnson Publications Ltd. 30s.
- Oldenburg, Henry (1967). *The Correspondence of Henry Oldenburg*, Vol. 4, 1667-1668. Ed. and trans. A. Rupert Hall & Marie Boas Hall. Madison, Wisconsin: University of Wisconsin Press. \$12.50. Pp. xxv+599.
- Oresme, Nicole (1968). *Maître Nicole Oresme, Le Livre du Ciel et du Monde*. Ed. Albert D. Menut & Alexander J. Denomy; trans. Albert D. Menut. Madison, Wisconsin: University of Wisconsin Press. \$17.50. Pp. xiii+778.

- Radnitzky, G. (1968). *Contemporary Schools of Metascience. I. Anglo-Saxon Schools of Metascience*. Göteborg, Sweden: Akademiförlaget Göteborg. n.p. Pp. xxiv+200.
- Radnitzky, G. (1968). *Contemporary Schools of Metascience. II. Continental Schools of Metascience*. Göteborg, Sweden: Akademiförlaget Göteborg. n.p. Pp. xii+200.
- de Santillana, G. (1968). *Reflections on Men and Ideas*. Cambridge, Mass.: M.I.T. Press. 140s. Pp. xiii+381.
- Shils, Edward, Ed. (1968). *Criteria for Scientific Development: Public Policy and National Goals*. A selection of articles from *Minerva*. Cambridge, Mass.: M.I.T. Press. \$8.95. Pp. xvi+207.
- Smith, Cyril Stanley, Ed. (1968). *Sources for the History of the Science of Steel 1532-1786*. Cambridge Mass.: The Society for the History of Technology with the M.I.T. Press. \$12.00. Pp. xv+357.
- Stegmüller, Wolfgang (1968). *Das Wahrheitsproblem und die Idee der Semantik: Eine Einführung in die Theorien von A. Tarski und R. Carnap*. Zweite, unveränderte Auflage. Vienna: Springer-Verlag [1st edn 1957]. DM 58.00. Pp. x+328.
- Tindal, Matthew (1967). *Christianity as old as the Creation*. Ed. G. Gawlick. Reprint edn. of the London edn. of 1730. Stuttgart: Friedrich Frommann Verlag. DM 64. Pp. 432.
- Vickers, Geoffrey (1968). *Value Systems and Social Process*. London: Tavistock Publications. 38s. Pp. xxii+217.
- Vico, Giambattista (1968). *The New Science of Giambattista Vico*. Trans. Thomas Goddard Bergin & Max Harold Fisch. Rev. trans. of 3rd edn (1744). Ithaca: Cornell University Press. \$15.00. Pp. xlv+441.
- Weil, Simone (1968). *On Science, Necessity and the Love of God*. London: Oxford University Press. 42s. Pp. x+20.

(b) ARTICLES

General

- Brownhill, R. S. (1968). Scientific ethics and the community. *Inquiry* 2, 243-8.
- Gotesky, R. (1968). The uses of inconsistency. *Phil. phenom. Res.* 28, 471-500.
- Malecki, Ignacy (1968). Science of science as a science and as practical activity. *Zagad. Nauk.* 4, 3-5.
- Marx, K. (1968). Issue devoted to Marxian thought. *Revta Filozof.* 15, 491-650.
- Szaniawski, Klemens (1968). Some remarks on the philosophy of science. *Zagad. Nauk.* 4, 26-32.
- Trasnea, Ovidiu (1968). Le concept de science politique aux XVI^e-XVII^e siècles. *Revta Filozof.* 15, 665-78.

Epistemology, Semantics, General Philosophical

- Apostol, Pavel (1968). Elemente pentru o epistemologie a stiintelor tehnice. *Revta Filozof.* 15, 427-36.
- Hanna, J. F. (1968). An explication of 'explication'. *Phil. Sci.* 35, 28-44.
- Joja, Crizantema (1965). Nature et objectifs de la philosophie des sciences. *Revue Roum. Scis socl* 9, 69-73.
- Long, D. C. (1968). Particulars and their qualities. *Phil. Q.* 18, 193-206.
- Margolis, J. (1968). Quine on observationality and translation. *Fdms Lang.* 4, 128-37.
- Resher, N. & Skyrms, F. B. (1968). A methodological problem in the evaluation of explanation. *Noûs* 2, 121-30.
- Wittgenstein, L. (1968). Notes for lectures on 'private experience' and 'sense data'. *Phil. Rev.* 77, 275-320.

Logic, Probability

- Beck, Heinrich (1968). Concepto y presupuestos gnoseológicos del método inductivo. *Anuario filosof.* 1, 13-22.
- Bohnert, H. G. (1967). Communication by Ramsey—sentence clause. *Phil. Sci.* 34, 341-7.
- Bolker, E. (1967). A simultaneous axiomatisation of utility and subjective probability. *Phil. Sci.* 34, 333-40.

- Cheng, Chung-Ying. (1968). Requirements for the validity of induction: an examination of Charles Peirce's theory. *Phil. phenom. Res.* 28, 392-402.
- Cohen, L. J. (1968). Geach's problem about intentional identity. *J. Phil.* 65, 329-34.
- Colby, K. M. & Enea, H. (1968). Inductive inference by intelligent machines. *Scientia* 103, 43-52.
- Dennett, D. C. (1968). Geach on intentional identity. *J. Phil.* 65, 335-40.
- Dumitriu, Anton (1967). Définition et existence. *Revue Roum. Scis socls* 11, 327-62.
- Dumitriu, Anton (1968). Le problème de paradoxes logico-mathématiques. *Scientia* 103, 394-408.
- Garner, R. T. (1968). Austin on entailment. *Phil. Q.* 18, 216-24.
- Hacking, I. (1967). Slightly more realistic personal probability. (Panel discussion with A. Shimony and L. J. Savage.) *Phil. Sci.* 34, 311-25.
- Harman, G. (1967). Detachment, probability and maximum likelihood. *Notls* 1, 401-11.
- MacQueen, James B. (1967). Some methods for classification and analysis of multivariate observations. *Proc 5th Berkeley Symp. math. Stat. Prob., 1965 and 1966*. University of California Press.
- Martin, R. L. (1968). On Grelling's paradox. *Phil. Rev.* 77, 321-31.
- Mihăilescu, Eugen (1965). Le problème de la décision dans la logique classique. *Revue Roum. Scis socls* 9, 51-67.
- Nakhnikian, G. (1968). In corrigibility. *Phil. Q.* 18, 207-15.
- Ono, Katuzi (1968). Does mathematics need something other than logic? *Ann. Japan Assn Phil. Sci.* 3, 7-18.
- Savage, L. J. (1967). Difficulties in the theory of personal probabilities. (Panel discussion with I. Hacking and A. Shimony.) *Phil. Sci.* 34, 305-10.
- Scientific hypotheses (1968). *J. Phil.* [Double issue] 10, 269-312.
- Shimony, A. (1967). Amplifying personal probability. (Panel discussion with L. J. Savage and I. Hacking.) *Phil. Sci.* 34, 326-32.
- Slupecki, Jerzy (1967). Deux interprétations du calcul à triple valeur des phrases. *Revue Roum. Scis socls* 11, 317-26.
- Turner, Merle B. (1968). Deciding for God—the Bayesian support of Pascal's wager. *Phil. phenom. Res.* 29, 84-90.
- Wertheimer, Roger (1968). Conditions. *J. Phil.* 65, 355-64.

Exact Sciences

- Bernadete, J. (1968). Continuity and the theory of measurement. *J. Phil.* 65, 411-30.
- Grünbaum, A. (1968). Are 'infinity machines' paradoxical? *Science* 159, 396-406.
- Hölling, J. (1968). Zur Kategorialanalyse des physikalischen Feldbegriffes. *Philosophia Naturalis* 10, 343-56.
- Hossfeld, P. (1968). Die Kategorie der Deszendenz bei N. Hartmann. *Philosophia Naturalis* 10, 333-42.
- Kar, Robert (1968). Die Spezielle Para-Relativitätstheorie. *Philosophia Naturalis* 10, 393-407.
- Koslow, A. (1968). Mach's concept of mass: program and definition. *Synthese* 18, 216-33.
- Margolis, J. (1968). Taylor on the reduction of teleological laws. *Inquiry* 2, 118-24. (See Taylor.)
- Rossteutscher, H. (1968). Über Kausalität und Komplementarität. *Philosophia Naturalis* 10, 356-68.
- Strobl, Wolfgang (1968). El principio de complementariedad y su significación científico-filosófica. *Anuario filosof.* 1, 185-206.
- Taylor, C. (1968). A reply to Margolis. *Inquiry* 2, 124-8. (See Margolis.)

Biological, Behavioural and Social Sciences

- Capek, M. (1968). Ernst Mach's biological theory of knowledge. *Synthese* 18, 171-91.
- Hall, T. S. (1968). On biological analogs of Newtonian paradigms. *Phil. Sci.* 35, 6-27.
- Stahl, H. H. (1968). Sociologie 'concreta' si istorie. *Revta Filozof.* 15, 379-93.
- Szarski, H. (1968). The concept of progress in evolution. *Scientia* 103, 152-60.

THE BRITISH JOURNAL FOR THE HISTORY OF SCIENCE

CONTENTS OF VOL. IV, PART II, No. 14

DECEMBER 1968

Main Articles:

Faraday, Matter and Natural Theology—Reflections on an Unpublished Manuscript

T. H. LEVERIE

The Geometrical Background to the "Merton School"

A. G. MOLLAND

John Mitchell and Henry Cavendish: Weighing the Stars

RUSSELL MCCORMMACH

Notes and Communications:

Galileo's Theory of Ballistics

PAUL-LAWRENCE ROSE

The Cook Bicentenary

D. W. WATERS

Essay Review

Book Reviews (22 books are reviewed)

PUBLISHED BY THE BRITISH SOCIETY FOR THE
HISTORY OF SCIENCE

Registered Office;

SCIENCE MUSEUM, SOUTH KENSINGTON, LONDON, S.W.7

THE AUSTRALASIAN JOURNAL OF PHILOSOPHY

EDITED BY GRAHAM NERLICH

VOL. 46

DECEMBER, 1968

No. 3

ARTICLES

CHARLES B. DANIELS

T' as a definite description

MENDEL F. COHEN

Wittgenstein's anti-essentialism

D. M. ARMSTRONG

The secondary qualities

HARRY S. SILVERSTEIN

Universalsability and egoism

C. A. HOOKER

Five arguments against Cragian transcriptionism

MAX DEUTSCHER

Popper's problem of an empirical base

REVIEWS

Published in May, August and December; Annual Subscription \$(Aus.) 4.50 (41/10 (Sterling)); \$(Aus.) 3.50 (32/6 (Sterling)) to members of the Australasian Association of Philosophy, which you are invited to join.

JOURNAL OF THE HISTORY OF IDEAS

A Quarterly devoted to Intellectual History

VOLUME XXIX

OCTOBER-DECEMBER 1968

NUMBER 4

TABLE OF CONTENTS

ARTICLES

- Homeric Psychology and the Oral Epic Tradition *Joseph A. Russo*
 Paradoxes of Solomon: Learning in the English Renaissance *M. Hattaway*
The Faerie Queene and Cosmogonies of the Near East *B. G. Feinsteln*
 "Interest Will Not Lie": A 17th-Century Political Maxim *J. A. W. Gunn*
 Galileo, Ficino, and More's *Psychathanasia* *C. A. Staudenbaur*
 Perry Miller's *New England Mind* *Gene Wise*

NOTES

- A Note on Stating the Socratic Paradox *John J. Mulhern*
 Buber's *Dialogue* and Gandhi's *Satyagraha* *V. V. Ramana Murthi*

REVIEW

- Encyclopedia of Philosophy*, edited by Paul Edwards Reviewed by *Phillip P. Wiener*

Subscription Rates (outside U.S.A. post paid): \$7.50 for one year;
 \$2.00 for single issue

Payable to

JOURNAL OF THE HISTORY OF IDEAS

The City College, New York, N.Y. 10031

SYNTHESIS

*An International Journal for Epistemology, Methodology and
 Philosophy of Science*

Editor: JAAKKO HINTIKKA

University of Helsinki and Stanford University

Board of Consulting Editors: David M. Armstrong, Yehoshua Bar-Hillel, Arthur W. Burks,
 Robert S. Cohen, Donald Davidson, Dagfinn Føllesdal, Hans Freudenthal, Béla Juhos,
 Jerrold J. Katz, Imre Lakatos, G. Nuchelmans, Wesley C. Salmon, Wolfgang Stegmüller,
 Erik Stenius, Patrick Suppes, Klemens Szaniawski, H. Törnebohm and Marx Wartofsky.

Contents of Volume 18, No. 4, October 1968

Including a Symposium on Thermodynamics

- Fred S. Roberts/R. Duncan Luce:* Axiomatic Thermodynamics and Extensive Measurement
J. J. Duistermaat: Energy and Entropy as Real Morphisms for Addition and Order
Fred Wilson: Is Operationism Unjust to Temperature?
Bengt Hansson: Fundamental Axioms for Preference Relations
Bengt Hansson: Choice Structures and Preference Relations
Carl R. Kordig: On Prescribing Description
Stephen Toulmin: Reply
Review: *Quantum Theory and Reality*, ed. by Mario Bunge (*Joseph D. Sneed*)
Books Received

Annual Subscription Price Dfl. 60.—(U.S.\$17.00)

Personal Subscription Price Dfl. 35.—(U.S.\$9.80)

D. REIDEL PUBLISHING COMPANY

DORDRECHT-HOLLAND

INQUIRY

An Interdisciplinary Journal of
Philosophy and the Social Sciences
Edited by ARNE NAESS

Volume 11

Number 4

Winter 1968

Historical Objectivity and Value Neutrality
The Covering Law Model of Historical Explanation
Situational Logic and Covering Law Explanation in History
Norm and Law in the Theory of Action
Discussion:

James J. Leach
Stanley Paluch
Michael Martin
Ruth Macklin

Wittgenstein and Self-Description

Vincent Hope

Review Discussion:

Arthur Danto: *Analytical Philosophy of History* and
Morton White: *Foundations of Historical Knowledge*

Leon Pompa

*Published quarterly by Universitetsforlaget, P.O. Box 307,
Blindern, Oslo 3, Norway*

Annual Subscription: 35s., \$6.00, postage included

Philosophy and Phenomenological Research

AN INTERNATIONAL QUARTERLY

Edited by MARVIN FARBER

DECEMBER, 1968, VOLUME XXIX, No. 2

CONTENTS

I. ARTICLES

Agency, by JOHN F. BOLER
Culture, Communication and Silence, by S. N. GANGULY
Husserl's Thought on God and Faith, by LOUIS DUPRÉ
The Neurotic as Moral Agent, by BEREL LANG
Features of Intentional Actions, by D. C. DENNETT
Are Persons Replaceable?, by KUANG-MING WU
Intentionality and Cognitive Sentences, by AUSONIO MARRAS
Cohen's Rationalistic Naturalism, by Y. H. KRIKORIAN

II. DISCUSSIONS

Josiah Royce and the American Race Problem, by WILLIAM T. FONTAINE
Professor Mackie on the Direction of Causation, by W. A. SUCHTING
Observations on a Programmed Course in Logic, by ROBERT E. GAHRINGER

III. REVIEWS, COMMUNICATIONS, NOTES AND NEWS, AND RECENT PUBLICATIONS

Subscription prices \$6.00 per year for individuals; \$7.50 per year for institutions.
Separate issues, \$1.65 and \$2.00 respectively.

Postage and handling for back issues only.

A circular listing the main contents of the Journal since 1940 will be sent upon request

Address all communications to

PHILOSOPHY AND PHENOMENOLOGICAL RESEARCH
State University of New York at Buffalo, Buffalo, New York, 14214

NOW AVAILABLE

Twenty year
Cumulative Index
1947-1967

THE REVIEW OF

METAPHYSICS

The Twenty year cumulative index covering the first twenty volumes of *The Review of Metaphysics* is divided into three sections: a subject index; an author index; and an index of books that have been examined in *Critical Studies*. \$3.00

Orders can now be placed for the Index.

For further details write to

THE REVIEW OF METAPHYSICS,
214 Lyman Beecher Hall, Haverford College,
Haverford, Pa. 19041.

MIND

A Quarterly Review of Psychology and Philosophy
edited by Professor GILBERT RYLE

JANUARY 1969

- | | |
|--|-------------------------------|
| I. On Self-Reference and a Puzzle in Constitutional Law | ALF ROSS |
| II. Epistemological Strategies | C. H. WHITELEY |
| III. Conventional Expressions of Emotion | LORENNE M. GORDON |
| IV. Forms as Causes in the <i>Phaedo</i> | C. C. W. TAYLOR |
| V. Thinking and Time | DAVID L. MOUTON |
| VI. Gödel and Physical Theory | WOLFGANG YOURGRAU |
| VII. Essential and Causal Explanations of Action | A. B. LEVISON and I. THALBERG |
| VIII. The Aesthetic Judgment and its Criteria of Value | SASCHA TALMOR |
| IX. Discussion Notes: | |
| Concerning 'Sense and Nonsense' | GORDON MATHESON |
| Neglecting to do what one can: A Reply | KEITH LEHRER |
| Some remarks on Act Utilitarianism | RICHARD T. GARNER |
| Zeno's Paradoxes and Continuity | IAN MUELLER |
| On Sommers' Logic of Sense and Nonsense | DAN FARSELL |
| Lehrer and Taylor on Modal Inference | LOWELL NISSEN |
| Rules and Consequences | J. BROOKS COLBURN |
| On the Unexpected Examination | A. K. AUSTIN |
| Linguistics and Ordinary Language | GRAHAM HUGH |
| A small point in Aristotle's Ethics | DAVID CODER |
| A Note on Black's "Companion" | BOGUSLAW WOLNIEWICZ |
| The Formalisation of Set Theory: A Reply to Mr. Swanson | JOHN TUCKER |
| X. Critical Notice | |
| Anthony Mannor: <i>Sartre: A Philosophic Study</i> | MARJORIE GRENE |
| XI. New Books | |
| XII. Notes | |
| XIII. List of Officers and Members of the Mind Association | |

Annual Subscription 25s, post free; single copy 7s 6d, plus postage

Published for MIND ASSOCIATION by

BASIL BLACKWELL

49 BROAD STREET OXFORD

d
i
a
l
e
c
t
i
c
a

dialectica

Revue trimestrielle de philosophie de la connaissance

- s'attache à promouvoir les principes directeurs de la *méthodologie ouverte*. Image vivante du développement de la philosophie ouverte, elle publie sur cette dernière et ses thèses essentielles des éditoriaux, des articles de fond, les comptes rendus des débats et des discussions sur ses tendances et ses engagements actuels
 - organise des débats et publie des enquêtes sur les questions essentielles de la connaissance moderne (voir par exemple les numéros sur l'idée de complémentarité avec la participation de cinq prix Nobel de physique, sur les fondements de la psychologie, du calcul des probabilités, sur la notion d'analogie, le fait et le droit, l'épistémologie de la physique quantique, etc.)
 - se prête à la remise en cause du fondement des disciplines abstraites et des bases des disciplines expérimentales
 - commente les publications récentes et accorde l'hospitalité de ses colonnes aux discussions de principe à leur sujet
 - appelle à la collaboration tous ceux qui entendent participer à la synthèse et à l'intégration de l'ensemble des connaissances modernes
 - vient de publier les résultats d'une enquête sur la question brûlante de la rénovation de l'enseignement des mathématiques, de la physique et de la biologie
 - prépare un symposium sur l'épistémologie de l'espace et du temps, poursuivra son enquête sur la méthode de la recherche, publiera prochainement des numéros sur l'état des recherches sur le fondement des mathématiques et sur la théorie générale du langage
- dialectica en un mot est l'organe vivant et la libre tribune de la *méthodologie* et de la *philosophie ouvertes*

Comité directeur: FERDINAND GONSETH, Lausanne

PAUL BERNAYS, Zurich

J.-B. GRIZE, Neuchâtel

P. NOLFI, Zurich

P. H. E. PILET, Lausanne

Rédaction: 12, chemin du Muveran, 1012 Lausanne (Suisse) Vente et abonnements: W. Roesch & Cie, Monbijoustrasse 9, 3000 Berne (Suisse)

Prix de vente au numéro: Suisse Fr 10 Etranger Fr 11
Abonnement annuel: Suisse Fr 35 Etranger Fr 40

Revue Internationale de Philosophie

Direction: Professeurs Ph. DEVAUX, H. GOUHIER, Ch. PERELMAN
Administration: Mme S. SERVAIS

99 avenue de l'Université, Bruxelles 5, Belgique.

Chaque numéro est consacré à l'étude d'un philosophe, d'un mouvement ou d'un problème philosophique, accompagnée souvent d'une bibliographie.

Vraiment internationale, la Revue recrute ses collaborateurs dans tous les pays et publie ses articles en allemand, anglais, espagnol, français et italien.

4 numéros par an

Abonnements: Pays européens: 400 FB. Autres pays: 450 FB.

Derniers numéros parus:

1967 n° 79/80: Josiah Royce (Bibliographie complète)

n° 81: Forme et pensée.

n° 82: La philosophie du langage. (Les précurseurs au XVIII^e siècle)

1968 n° 83/84: Eugène Dupréel.

A paraître:

La crise de l'humanisme dans la pensée contemporaine.

Wittgenstein.

L'Analogie.

T H E O R I A

A Swedish Journal of Philosophy

VOLUME XXXIV

1968

PART 1

CONTENTS

Stig Kanger: Equivalent theories

Krister Segerberg: Decidability of S4.1

Krister Segerberg: Decidability of four modal logics

Krister Segerberg: Propositional logics related to Heyting's and Johansson's

Jullan Wolfe: The criteria of sleep

Ardon Lyon: Family resemblance, vagueness, and change of meaning

Dag Prawitz: A discussion note on utilitarianism

Robert Binkley and Romane Clark: A cancellation algorithm corrected

CWK Glerup, Öresundsvägen 1, Lund, Sweden

FOUNDATIONS OF LANGUAGE

International Journal of Language and Philosophy

Editorial Board: Morris Halle (MIT), Peter Hartmann (Münster/W), K. Kunjuni Raja (Madras), Benson Mates (Univ. of California), J. F. Staal (Univ. of California), Pieter A. Verburg (Groningen), John W. M. Verhaar (Nijmegen)

VOLUME 4, No. 3, AUGUST 1968

Contents: H. P. Grice: Utterer's Meaning, Sentence-Meaning, and Word-Meaning—James D. McCawley: Concerning the Base Component of a Transformational Grammar—Roland Harweg: Language and Music—An Immanent and Sign Theoretic Approach: Some Preliminary Remarks—Chung-Ying Cheng: Eliminability of Singular Terms Reconsidered—A. R. Luria and L. S. Trvetkova: The Mechanism of "Dynamic Aphasia"—John Anderson: On the Status of "Lexical Formatives"—Marshall Durbin and Michael Micklin: Sociolinguistics: Some Methodological Contributions from Linguistics—Reviews by William Kneale and James Deese—Publications Received for Review.

FOUNDATIONS OF LANGUAGE is published quarterly

Annual subscription price Dfl.42.—(US \$11.75)

D. REIDEL PUBLISHING COMPANY/DORDRECHT - HOLLAND

Studium Generale

Journal for Interdisciplinary Studies

Executive Editor:
G. H. Müller (Heidelberg)

Reviews:
G. G. Grau (Hannover)

STUDIUM GENERALE has been principally concerned ever since its foundation to clarify relations between the various branches of science, endeavouring in this way to provide the materials for a new image of World and Man.

STUDIUM GENERALE now publishes presentations of new research results in science and the humanities; studies in the history of problems and concepts; epistemological studies (philosophy of science); well-founded hypotheses predicting the advance of science, technology and social behaviour (forecasting the future); and attempts to integrate the image of World and Man.

STUDIUM GENERALE will call increasingly upon English-speaking authors so as to encourage the exchange of seminal ideas about the direction and consequences of research between Europe and the Western hemisphere.

STUDIUM GENERALE also contains a review section giving short summaries of articles in the leading journals on philosophy and the philosophy of science.

General Subjects 1968/69: American Literature. Dialectics. Government. Cybernetics. Magnitudes of Universe. Linguistics. Study of Time. Medical, Legal and Ethical Problems Thrown up by Heart Transplants. Philosophy of Politics. Reviews.

STUDIUM GENERALE is published monthly; 12 issues, each of approximately 100 pages, constitute one volume.
Subscription price Vol. 22, 1969: US \$23.70 including postage. (Subscriptions are entered on a prepaid basis only).
Single copies US \$2.25

Please contact your bookstore for descriptive leaflet.

Students, research fellows and assistant lecturers can obtain STUDIUM GENERALE on subscription at 20% reduction.



Springer - Verlag
Berlin · Heidelberg · New York

PHILOSOPHIA NATURALIS

*Archiv für Naturphilosophie und die philosophischen Grenzgebiete
der exakten Wissenschaften und Wissenschaftsgeschichte*

Begründet von EDUARD MAY†

Herausgegeben von JOSEPH MEURERS

Inhalt Heft 4, Band 1°

M. BORN und W. BIEM: *Dualismus in der Quantentheorie*

B. KANITSCHNEIDER: *Zur Frage der Naturbeschreibung*

B. J. VAN EYK: *Über die Begrenzung der Anwendbarkeit klassischer wissenschaftliche
Konzeptionen auf dem Gebiet der Naturwissenschaften*

P. MITTELSTAEDT: *Verborgene Parameter und beobachtbare Größen in physikalischen
Theorien*

ZEITSCHRIFTENSCHAU VON P. HOSSFELD

Vier Hefte bilden einen Band: 40 DM

Umfang eines Bandes: 512 Seiten

Einzelheft, soweit verfügbar: 11 DM

VERLAG ANTON HAIN KG . 6554 MEISENHEIM
(West-Germany)



“SCIENTIA”

INTERNATIONAL REVIEW OF SCIENTIFIC SYNTHESIS

(1969 - 63rd Year)

Scientific Committee

G. ABETTI - G. B. BONINO - P. CALDIROLA - L. CALIFANO - R. CANESTRARI -
G. COLONNETTI - B. DE FINETTI - G. MERLA - G. MONTALENTI - E. PADOA - G. SANSONE
A. SESTINI - V. SOMENZI - S. TONZIG.

Editor N. BONETTI

“SCIENTIA” publishes original articles by leading workers in all the fields
of science and philosophy, from all countries.

It is a magazine of synthesis, that is, of the convergence of the various branches
of knowledge.

It is a magazine of multidisciplinary information; it promotes the “bridge”
disciplines, which are the most active in the progress of research.

It encourages the building of a culture fully adherent to modern scientific and
philosophic standards.

Published every two months, each issue containing approximately 200 pages.
Articles are published in the original language of the author i.e. English,
Italian, French, Spanish and German. Every issue contains a supplement
with a French translation of the non-French articles. Each issue contains:
Book reviews - Analysis - Chronicle.

SUBSCRIPTION RATES: Italy Lire 15.000 - Abroad \$32.00 or £13
Current issue: Italy Lire 3.000 - Abroad \$6.50 or £2.

Back issues: Italy Lire 500 - Abroad \$1.50 or 11s. (in stamps).

“Scientia” representative in U.K.: Atlas Publishing & Distributing Co. Ltd. -
334 Brixton Road - London S.W.9

“SCIENTIA” - Via Donizetti, 55 - 20122 Milano (Italy)

THE BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

Committee 1968-69

President: Mr J. L. Mackie

Vice-President: Dr G. J. Whitrow

Hon. Secretary: Mr R. Harré

Hon. Treasurer: Dr W. Mays

Hon. Secretary of the Northern Group: Dr N. Wetherick

Ordinary Members: Dr J. Giedymin, Mr A. E. Musgrave, Mr H. D. Sluga,
Dr Eva Cassirer

EDITORIAL BOARD OF THE JOURNAL

Professor J. D. Bernal

Professor R. B. Braithwaite

Professor H. Dingle

Professor Sir John Eccles

Professor Sir Harold Jeffreys

Professor W. C. Kneale

Professor S. Körner

Professor M. Polanyi

Professor Sir Karl Popper

Earl Russell

Dr G. J. Whitrow

Mr L. L. Whyte

Professor J. H. Woodger

The British Journal for the Philosophy of Science is the official organ of the Society, whose purpose is described by the following statement from its Constitution and Rules :

‘ The purpose of the Society is to study the logic, the method and the philosophy of science as well as those of the various special sciences, including the social sciences.’

The Society holds meetings, at which papers are read and discussed, once a month during the academic year. These are held, by kind permission, in University College, London.

Ordinary membership is open to all persons who have been proposed by two members, approved by the Committee, and elected at an ordinary meeting. The annual subscription is 20s. Those who wish to join should write to the Hon. Secretary at Linacre College, Oxford.

The Northern Group of the Society meets regularly at various centres in the north of England. Those who wish to join it should write to the Hon. Secretary of the Northern Group at the Undergraduate School of Studies in the Social Sciences, University of Bradford, Bradford 7.



Brit. J. Phil. Sci. Vol. 19 No. 4 February 1969

The British Journal for the Philosophy of Science

Articles

- 281-299 R. G. SWINBURNE Vagueness, Inexactness, and Imprecision
301-316 RUTH MACKLIN Action, Causality, and Teleology
317-324 H. A. C. DOBBS The 'Present' in Physics
325-344 P. M. WILLIAMS The Structure of Acceptance and its
Evidential Basis

Discussions

- 345-353 WILLIAM W. ROZEBOOM New Mysteries for Old: the
Transfiguration of Miller's Paradox
353-356 DHARMENDRA KUMAR Neutrality, Contingency, and Un-
decidability
357-358 I. J. GOOD Gödel's Theorem is a Red Herring

Reviews

Recent Publications

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, N.W.1.

American Branch: 32 East 57th Street, New York, N.Y. 10022

Single issues 20s. net in U.K. (\$2.75 in U.S.A.)

Subscription price 60s. net in U.K. (\$9.50 in U.S.A.)

3 APR 1969